

**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)



16 SEP 2003

REC'D 26 SEP 2003

WIPO

PCT

**Prioritätsbescheinigung über die Einreichung
einer Patentanmeldung**

Aktenzeichen:

102 38 978.0

Anmeldetag:

20. August 2002

Anmelder/Inhaber:

SunGene GmbH & Co. KGaA, Gatersleben/DE

Bezeichnung:

Verfahren zur Herstellung von Ketocarotinoiden in
Früchten von Pflanzen

IPC:

A 01 H, C 12 N, A 23 K

**Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ur-
sprünglichen Unterlagen dieser Patentanmeldung.**

München, den 9. September 2003
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Stanschus

Patentansprüche

1. Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von genetisch veränderten Pflanzen, die in Früchten eine Ketolase-Aktivität aufweisen.
5
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass man genetisch veränderte Pflanzen verwendet, die in Früchten eine Ketolase exprimieren.
10
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass man genetisch veränderte Pflanzen verwendet, die in Früchten mindestens eine Nukleinsäure, kodierend eine Ketolase, enthalten.
15
4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, dass man genetisch veränderte Pflanzen verwendet, in die man ausgehend von einer Ausgangspflanze mindestens eine Nukleinsäure, kodierend eine Ketolase, eingebracht hat.
20
5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, dass man Nukleinsäuren einbringt, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20% auf Aminosäureebene mit der Sequenz SEQ ID NO. 2 und die enzymatische Eigenschaft einer Ketolase aufweist.
25
6. Verfahren nach Anspruch 4, dadurch gekennzeichnet, dass man Nukleinsäuren, enthaltend die Sequenz SEQ ID NO. 1 einbringt.
30
7. Verfahren nach Anspruch 4, dadurch gekennzeichnet, dass man Nukleinsäuren einbringt die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO. 16 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20% auf Aminosäureebene mit der Sequenz SEQ ID NO. 16 und die enzymatische Eigenschaft einer Ketolase aufweist.
35
8. Verfahren nach Anspruch 6, dadurch gekennzeichnet, dass man Nukleinsäuren, enthaltend die Sequenz SEQ ID NO. 15 einbringt.
40

45

2

9. Verfahren nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, dass man genetisch veränderte Pflanzen verwendet, die in Früchten die höchste Expressionsrate einer Ketolase aufweisen.
- 5
10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, dass die Genexpression der Ketolase unter Kontrolle eines fruchtspezifischen Promotors erfolgt.
- 10 11. Verfahren nach einem der Ansprüche 1 bis 10, dadurch gekennzeichnet, dass man als Pflanze eine Pflanze verwendet, die in Früchten Chromoplasten aufweist.
12. Verfahren nach einem der Ansprüche 1 bis 11, dadurch gekennzeichnet, dass man als Pflanze eine Pflanze ausgewählt aus den Pflanzengattungen Actinophloeus, Aglaeonema, Ananas, Arbutus, Archontophoenix, Area, Aronia, Asparagus, Attalea, Berberis, Bixia, Brachychilum, Bryonia, Caliptocalix, Capsicum, Carica, Celastrus, Citrullus, Citrus, Convallaria, Cotoneaster, Crataegus, Cucumis, Cucurbita, Cuscuta, Cycas, Cyphomandra, Dioscorea, Diospyrus, Dura, Elaeagnus, Elaeis, Erythroxylon, Euonymus, Ficus, Fortunella, Fragaria, Gárdinia, Gonocaryum, Gossypium, Guava, Guilielma, Hibiscus, Hippophaea, Iris, Lathyrus, Lonicera, Luffa, Lycium, Lycopersicum, Malpighia, Mangifera, Mormodica, Murraya, Musa, Nenga, Palisota, Pandanus, Passiflora, Persea, Physalis, Prunus, Ptychandra, Punica, Pyracantha, Pyrus, Ribes, Rosa, Rubus, Sabal, Sambucus, Seaforita, Shepherdia, Solanum, Sorbus, Synaspadix, Tabernae, Tamus, Taxus, Trichosanthes, Triphasia, Vaccinium, Viburnum, Vignia oder Vitis verwendet.
- 15
- 20
- 25
- 30
13. Verfahren nach einem der Ansprüche 1 bis 12, dadurch gekennzeichnet, dass man nach dem Kultivieren die genetisch veränderten Pflanzen erntet und anschließend die Ketocarotinoide aus den Früchten der Pflanzen isoliert.
- 35
14. Verfahren nach einem der Ansprüche 1 bis 13, dadurch gekennzeichnet, dass die Ketocarotinoide ausgewählt sind aus der Gruppe Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin und Adonixanthin.
- 40
15. Nukleinsäurekonstrukt, enthaltend funktionell verknüpft einen fruchtspezifischen Promotor und eine Nukleinsäure kodierend eine Ketolase.
- 45

3

16. Genetisch veränderte Pflanze, die in Früchten eine Ketolase-Aktivität aufweist.
17. Genetisch veränderte Pflanze nach Anspruch 16, dadurch gekennzeichnet, dass die genetisch veränderte Pflanze in den Früchten eine Ketolase exprimiert.
18. Genetisch veränderte Pflanze nach Anspruch 16 oder 17, enthaltend in Früchten mindestens eine Nukleinsäure, kodierend eine Ketolase.
19. Genetisch veränderte Pflanze nach einem der Ansprüche 16 bis 18, dadurch gekennzeichnet dass man in die Pflanze ausgehend von einer Ausgangspflanze mindestens eine Nukleinsäure, kodierend eine Ketolase, eingebracht hat.
20. Genetisch veränderte Pflanze, ausgewählt aus den Pflanzengattungen Actinophloeus, Aglaeonema, Ananas, Arbutus, Archontophoenix, Area, Aronia, Asparagus, Attalea, Berberis, Bixia, Brachychilum, Bryonia, Caliptocalix, Capsicum, Carica, Celastrus, Citrullus, Citrus, Convallaria, Cotoneaster, Crataegus, Cucumis, Cucurbita, Cuscuta, Cycas, Cyphomandra, Dioscorea, Diospyrus, Dura, Elaeagnus, Elaeis, Erythroxylon, Euonymus, Ficus, Fortunella, Fragaria, Gardinia, Gonocaryum, Gossypium, Guava, Guilielma, Hibiscus, Hippophaea, Iris, Lathyrus, Lonicera, Luffa, Lycium, Lycopersicum, Malpighia, Mangifera, Mormodica, Murraya, Musa, Nenga, Palisota, Pandanus, Passiflora, Persea, Physalis, Prunus, Ptychandra, Punica, Pyracantha, Pyrus, Ribes, Rosa, Rubus, Sabal, Sambucus, Seaforita, Shepherdia, Solanum, Sorbus, Synaspadix, Tabernae, Tamus, Taxus, Trichosanthes, Triphasia, Vaccinium, Viburnum, Vignia oder Vitis, enthaltend mindestens eine Nukleinsäure, kodierend eine Ketolase.
21. Genetisch veränderte Pflanze nach Anspruch 20, dadurch gekennzeichnet, dass die Ketolase in Früchten exprimiert wird.
22. Genetisch veränderte Pflanze nach einem der Ansprüche 16 bis 21, dadurch gekennzeichnet, dass die Expressionsrate einer Ketolase in Früchten am höchsten ist.
23. Verwendung der genetisch veränderten Pflanzen nach einem der Ansprüche 16 bis 22 als Futter- oder Nahrungsmittel.

4

24. Verwendung der Früchte der genetisch veränderten Pflanzen nach einem der Ansprüche 17 bis 23 zur Herstellung von Keto-carotinoid-haltigen Extrakten oder zur Herstellung von Futter- oder Nahrungsergänzungsmittel.

5

25. Verfahren zur Herstellung von genetisch veränderten Pflanzen gemäß Anspruch 22, dadurch gekennzeichnet, dass man ein Nukleinsäurekonstrukt, enthaltend funktionell verknüpft einen fruchtspezifischen Promotor und Nukleinsäuren kodierend

10

eine Ketolase in das Genom der Ausgangspflanze einführt.

15

20

25

30

35

40

45

Verfahren zur Herstellung von Ketocarotinoiden in Früchten von Pflanzen

5 Beschreibung

Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von genetisch veränderten Pflanzen, die in Früchten eine Ketolase-Aktivität aufweisen, die-
10 genetisch veränderten Pflanzen, sowie deren Verwendung als Nahrungs- oder Futtermittel und zur Herstellung von Ketocarotinoidextrakten.

Carotinoide werden de novo in Bakterien, Algen, Pilzen und Pflanzen synthetisiert. Ketocarotinoide, also Carotinoide, die mindestens eine Keto-Gruppe enthalten, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin oder Adonixanthin sind natürliche Antioxidantien und Pigmente, die von einigen Algen und Mikroorganismen
20 als Sekundärmetabolite produziert werden.

Aufgrund ihrer farbgebenden Eigenschaften werden die Ketocarotinoide und insbesondere Astaxanthin als Pigmentierhilfsstoffe in der Tierernährung, insbesondere in der Forellen-, Lachs- und
25 Shrimpszucht verwendet.

Die Herstellung von Astaxanthin erfolgt heutzutage größtenteils durch chemische Syntheseverfahren. Natürliche Ketocarotinoide, wie beispielsweise natürliches Astaxanthin, werden heutzutage in
30 biotechnologischen Verfahren in kleinen Mengen durch Kultivierung von Algen, beispielsweise Haematococcus pluvialis, oder durch Fermentation von gentechnologisch optimierten Mikroorganismen und anschließender Isolierung gewonnen.

35 Ein wirtschaftliches biotechnologisches Verfahren zur Herstellung von natürlichen Ketocarotinoiden ist daher von großer Bedeutung.

WO 98/18910 beschreibt die Synthese von Ketocarotinoiden in Nektarien von Tabakblüten durch Einbringen eines Ketolase-Gens
40 in Tabak.

WO 01/20011 beschreibt ein DNA Konstrukt zur Produktion von Ketocarotinoiden, insbesondere Astaxanthin, in Samen von Ölsaatpflanzen wie Raps, Sonnenblume, Sojabohne und Senf unter Verwendung
45 eines Samen-spezifischen Promotors und einer Ketolase aus Haematococcus.

2

Die im Stand der Technik offenbarten Verfahren liefern zwar genetisch veränderte Pflanzen, die in spezifischen Geweben einen Gehalt an Ketocarotinoiden aufweisen, weisen jedoch den Nachteil auf, dass die Höhe des Gehalts an Ketocarotinoiden und die Reinheit, insbesondere an Astaxanthin, noch nicht zufriedenstellend ist.

Der Erfindung lag daher die Aufgabe zugrunde, ein alternatives Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von Pflanzen zur Verfügung zu stellen, bzw. weitere transgene Pflanzen, die Ketocarotinoide herstellen, zur Verfügung zu stellen, die optimierte Eigenschaften, wie beispielsweise einen höheren Gehalt an Ketocarotinoiden, aufweisen und den geschilderten Nachteil des Standes der Technik nicht aufweisen.

15

Demgemäß wurde ein Verfahren zur Herstellung von Ketocarotinoiden gefunden, indem man genetisch veränderte Pflanzen kultiviert, die in Früchten eine Ketolase-Aktivität aufweisen.

20 Unter Ketolase-Aktivität wird die Enzymaktivität einer Ketolase verstanden.

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

25

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Canthaxanthin umzuwandeln.

30

Dementsprechend wird unter Ketolase-Aktivität die in einer bestimmten Zeit durch das Protein Ketolase umgesetzte Menge β -Carotin bzw. gebildete Menge Canthaxanthin verstanden.

35 Um in den Früchten der genetisch veränderten Pflanzen eine Ketolaseaktivität aufzuweisen, werden in einer bevorzugten Ausführungsform genetisch veränderte Pflanzen verwendet, die in Früchten eine Ketolase exprimieren.

40 Vorzugsweise werden daher im erfindungsgemäßen Verfahren genetisch veränderte Pflanzen verwendet, die in Früchten mindestens eine Nukleinsäure, kodierend eine Ketolase, enthalten.

45

3

Es sind keine Pflanzen bekannt, die als Wildtyp in Früchten eine Ketolase-Aktivität aufweisen. Insbesondere weisen die nachstehend beschriebenen, bevorzugten Pflanzen in Früchten als Wildtyp keine Ketolase-Aktivität auf.

5

In der vorliegenden Erfindung wird die Ketolase-Aktivität in Früchten der genetisch veränderten Pflanzen durch die genetische Veränderung der Ausgangspflanze verursacht. Die erfindungsgemäße genetisch veränderte Pflanze weist somit, im Vergleich

10 zur genetisch nicht veränderten Ausgangspflanze eine Ketolase-Aktivität in Früchten auf und ist somit vorzugsweise in der Lage, in Früchten eine Ketolase zu exprimieren.

Unter dem Begriff "Ausgangspflanze" oder "Wildtyp" wird die ent-
15 sprechende nicht genetisch veränderte Ausgangspflanze verstanden.

Unter dem Begriff "genetisch veränderte Pflanze" wird vorzugsweise eine im Vergleich zur Ausgangspflanze genetisch veränderte Pflanze verstanden.

20

Je nach Zusammenhang kann unter dem Begriff "Pflanze" die Ausgangspflanze (Wildtyp) oder eine erfindungsgemäße, genetisch veränderte Pflanze oder beides verstanden werden.

25 Die Verursachung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, in den Früchten der Pflanzen erfolgt vorzugsweise durch Einbringen von Nukleinsäuren, die Ketolasen kodieren, in die Ausgangspflanze.

30 Die Erfindung betrifft daher insbesondere das vorstehend beschriebene Verfahren, dadurch gekennzeichnet, dass man genetisch veränderte Pflanzen verwendet, in die man ausgehend von einer Ausgangspflanze, mindestens eine Nukleinsäure, kodierend eine Ketolase, eingebracht hat.

35

Dazu kann prinzipiell jedes Ketolase-Gen, also jede Nukleinsäure die eine Ketolase kodiert, verwendet werden.

Alle in der Beschreibung erwähnten Nukleinsäuren können
40 beispielsweise eine RNA-, DNA- oder cDNA-Sequenz sein.

Bei genomischen Ketolase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall, dass die Wirtspflanze nicht in der Lage ist oder nicht in die Lage versetzt werden

45 kann, die entsprechenden Ketolase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs, zu verwenden.

4

Beispiele für Nukleinsäuren, kodierend eine Ketolase, und die entsprechenden Ketolasen, die im erfindungsgemäßen Verfahren bzw. in den nachstehend beschriebenen erfindungsgemäßen genetisch veränderten Pflanzen verwendet werden können, sind beispielsweise

5 Sequenzen aus

Haematoccus pluvialis, insbesondere aus Haematoccus pluvialis Flotow em. Wille (Accession No. X86782; Nukleinsäure: SEQ ID No. 1, Protein SEQ ID No. 2),

10

Haematoccus pluvialis, NIES-144 (Accession No. D45881; Nukleinsäure: SEQ ID No. 3, Protein SEQ ID No. 4),

Agrobacterium aurantiacum (Accession No. D58420; Nukleinsäure:

15 SEQ. ID. No. 5, Protein SEQ ID No. 6),

Aliccaligenes spec. (Accession No. D58422; Nukleinsäure: SEQ ID No. 7, Protein SEQ ID No. 8),

20 Paracoccus marcusii (Accession No. Y15112; Nukleinsäure: SEQ ID No. 9, Protein SEQ ID No. 10).

Synechocystis sp. Strain PC6803 (Accession No. S76617, NP442491; Nukleinsäure: SEQ ID No. 11, Protein SEQ ID No. 12).

25

Bradyrhizobium sp. (Accession No. AF218415, BAB 74888; Nukleinsäure: SEQ ID No. 13, Protein SEQ ID No. 14).

Nostoc sp. Strain PCC7120 (Accession No. AP003592; Nukleinsäure:

30 SEQ ID No. 15, Protein SEQ ID No. 16).

Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene, die im erfindungsgemäßen Verfahren verwendet werden können, lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen und insbesondere mit den Sequenzen SEQ ID NO. 2 und/oder SEQ ID NO. 16 leicht auffinden.

40

Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID. No 1 und/oder SEQ ID NO. 15 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

45

5

Die Hybridisierung kann unter moderaten (geringe Stringenz) oder vorzugsweise unter stringenten (hohe Stringenz) Bedingungen erfolgen.

- 5 Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben.

- Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit 2X SSC bei 50°C) und solchen mit hoher Stringenz (mit 0.2X SSC bei 50°C, bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M Natriumchlorid, pH 7.0).

- Darüberhinaus kann die Temperatur während des Waschschrilles von moderaten Bedingungen bei Raumtemperatur, 22°C, bis zu stringenten Bedingungen bei 65°C angehoben werden.

- Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50 % Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt.

- Einige beispielhafte Bedingungen für Hybridisierung und Waschschrill sind infolge gegeben:

(1) Hybridisierungsbedingungen mit zum Beispiel

- (i) 4X SSC bei 65°C, oder
- (ii) 6X SSC bei 45°C, oder
- (iii) 6X SSC bei 68°C, 100 mg/ml denaturierter Fischsperma-DNA, oder
- (iv) 6X SSC, 0,5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA bei 68°C, oder
- (v) 6XSSC, 0,5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA, 50 % Formamid bei 42°C, oder

6

- (vi) 50 % Formamid, 4X SSC bei 42°C, oder
- (vii) 50 % (vol/vol) Formamid, 0,1 % Rinderserumalbumin, 0,1 % Ficoll, 0,1 % Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6,5, 750 mM NaCl, 75 mM Natriumcitrat bei 42°C, oder
- (viii) 2X oder 4X SSC bei 50°C (moderate Bedingungen), oder
- (ix) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42° (moderate Bedingungen).
- (2) Waschschritte für jeweils 10 Minuten mit zum Beispiel
- (i) 0,015 M NaCl/0,0015 M Natriumcitrat/0,1 % SDS bei 50°C, oder
- (ii) 0,1X SSC bei 65°C, oder
- (iii) 0,1X SSC, 0.5 % SDS bei 68°C, oder
- (iv) 0,1X SSC, 0.5 % SDS, 50 % Formamid bei 42°C, oder
- (v) 0,2X SSC, 0.1 % SDS bei 42°C, oder
- (vi) 2X SSC bei 65°C (moderate Bedingungen).

In einer bevorzugten Ausführungsform der erfindungsgemäßen Verfahrens bringt man Nukleinsäuren ein, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO. 2 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 % auf Aminosäureebene mit der Sequenz SEQ ID NO. 2 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie vorstehend beschrieben durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO. 2 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

7

In einer weiteren bevorzugten Ausführungsform der erfindungsgemäßen Verfahren bringt man man Nukleinsäuren ein, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO. 16 oder eine von dieser Sequenz durch Substitution, Insertion oder

5 Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 % auf Aminosäureebene mit

10 der Sequenz SEQ ID NO. 16 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie vorstehend beschrieben durch Identitätsvergleich der

15 Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO. 16 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

20 Unter dem Begriff "Substitution" ist in der Beschreibung der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren zu verstehen. Bevorzugt werden sog. konservative Austausche durchgeführt, bei denen die ersetzte Aminosäure eine ähn-

25 liche Eigenschaft hat wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

Deletion ist das Ersetzen einer Aminosäure durch eine direkte

30 Bindung. Bevorzugte Positionen für Deletionen sind die Termini des Polypeptides und die Verknüpfungen zwischen den einzelnen Proteindomänen.

Insertionen sind Einfügungen von Aminosäuren in die Polypeptid-

35 kette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren ersetzt wird.

Unter Identität zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinlänge verstanden,

40 insbesondere die Identität die durch Vergleich mit Hilfe der Lasergene Software der Firma DNASTAR, inc. Madison, Wisconsin (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung

45 folgender Parameter berechnet wird:

8

Multiple alignment parameter:

Gap penalty	10
Gap length penalty	10

5 Pairwise alignment parameter:

K-tuple	1
Gap penalty	3
Window	5
Diagonals saved	5

10

Unter einem Protein, das eine Identität von mindestens 20 % auf Aminosäureebene mit der Sequenz SEQ ID NO. 2 oder 16 aufweist, wird dementsprechend ein Protein verstanden, das bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO. 2 oder 16, insbesondere nach obigen Programmalgorithmus mit obigem Parametersatz eine Identität von mindestens 20 % aufweist.

15

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

20

Bevorzugt werden dafür solche Kodons verwendet, die entsprechend der pflanzespezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

25

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO. 1, in den Pflanze ein.

30

In einer weiteren besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO. 15, in den Pflanze ein.

35 Alle vorstehend erwähnten Ketolase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning:

45

A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

In einer besonderes bevorzugten Ausführungsform der erfindungs-
5 gemäßen Verfahrens verwendet man genetisch veränderte Pflanzen, die in Früchten die höchste Expressionsrate einer Ketolase aufweisen.

Vorzugsweise wird dies dadurch erreicht, dass die Genexpression
10 der Ketolase unter Kontrolle eines fruchtspezifischen Promotors erfolgt. Beispielsweise werden dazu die vorstehend beschriebenen Nukleinsäuren, wie nachstehend ausführlich beschrieben, in einem Nukleinsäurekonstrukt, funktionell verknüpft mit einem fruchtspezifischen Promotor, in die Pflanze eingebracht.

15 Unter Pflanzen werden erfindungsgemäß vorzugsweise Pflanzen verstanden, die als Wildtyp in Früchten Chromoplasten aufweisen.

Weiter bevorzugte Pflanzen weisen als Wildtyp in den Früchten zu-
20 sätzlich Carotinoide, insbesondere β -Carotin, Zeaxanthin, Neoxanthin, Violaxanthin oder Lutein auf.

Weiter bevorzugte Pflanzen weisen als Wildtyp in den Früchten zu-
sätzlich eine Hydroxylase-Aktivität auf.

25 Unter Hydroxylase-Aktivität wird die Enzymaktivität einer Hydroxylase verstanden.

Unter einer Hydroxylase wird ein Protein verstanden, das die
30 enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

Inbesondere wird unter einer Hydroxylase ein Protein verstanden,
35 das die enzymatische Aktivität aufweist, β -Carotin in Zeaxanthin oder Canthaxanthin in Astaxanthin umzuwandeln.

Dementsprechend wird unter Hydroxylase-Aktivität die in einer
bestimmten Zeit durch das Protein Hydroxylase umgesetzte Menge
40 β -Carotin oder Canthaxanthin bzw. gebildete Menge Zeaxanthin oder Astaxanthin verstanden.

Besonders bevorzugte Pflanzen sind Pflanzen, ausgewählt aus den
Pflanzengattungen Actinophloeus, Aglaeonema, Ananas, Arbutus,
45 Archontophoenix, Area, Aronia, Asparagus, Attalea, Berberis, Bixia, Brachychilum, Bryonia, Caliptocalix, Capsicum, Carica, Celastrus, Citrullus, Citrus, Convallaria, Cotoneaster,

10

Crataegus, Cucumis, Cucurbita, Cuscuta, Cycas, Cyphomandra, Dioscorea, Diospyrus, Dura, Elaeagnus, Elaeis, Erythroxylon, Euonymus, Ficus, Fortunella, Fragaria, Gardinia, Gonocaryum, Gossypium, Guava, Guilielma, Hibiscus, Hippophaea, Iris, Lathyrus, 5 Lonicera, Luffa, Lycium, Lycopersicum, Malpighia, Mangifera, Mormodica, Murraya, Musa, Nenga, Palisota, Pandanus, Passiflora, Persea, Physalis, Prunus, Ptychandra, Punica, Pyracantha, Pyrus, Ribes, Rosa, Rubus, Sabal, Sambucus, Seaforita, Shepherdia, Solanum, Sorbus, Synaspadix, Tabernae, Tamus, Taxus, Trichosanthes, 10 Triphasia, Vaccinium, Viburnum, Vignia oder Vitis.

Die Bestimmung der Ketolase-Aktivität in erfindungsgemäßen genetisch veränderten Pflanzen erfolgt in Anlehnung an die Methode von Frazer et al., (J. Biol. Chem. 272(10): 6128-6135, 1997). Die 15 Ketolase-Aktivität in pflanzlichen Extrakten wird mit den Substraten beta-Carotin und Canthaxanthin in Gegenwart von Lipid (Sojalecithin) und Detergens (Natriumcholat) bestimmt. Substrat/Produkt-Verhältnisse aus den Ketolase-Assays werden mittels HPLC ermittelt.

20

Im erfindungsgemäßen Verfahren zur Herstellung von Ketocarotinoiden wird vorzugsweise dem Kultivierungsschritt der genetisch veränderten Pflanzen, im folgenden auch transgene Pflanzen bezeichnet, ein Ernten der Pflanzen und ein Isolieren von Ketocarotinoiden 25 aus den Früchten der Pflanzen angeschlossen.

Die transgenen Pflanzen werden in an sich bekannter Weise auf Nährböden gezogen und entsprechend geerntet.

30 Die Isolierung von Ketocarotinoiden aus den geernteten Früchten erfolgt in an sich bekannter Weise, beispielsweise durch Trocknung und anschließender Extraktion und gegebenenfalls weiterer chemischer oder physikalischer Reinigungsprozesse, wie beispielsweise Fällungsmethoden, Kristallographie, thermische Trennverfahren, wie Rektifizierverfahren oder physikalische Trennverfahren, wie beispielsweise Chromatographie. Die Isolierung von 35 Ketocarotinoiden aus den Früchten erfolgt beispielsweise bevorzugt durch organische Lösungsmittel wie Aceton, Hexan, Ether oder tert.-Methylbutylether.

40

Weitere Isolierverfahren von Ketocarotinoiden sind beispielsweise in Egger und Kleinig (Phytochemistry (1967) 6, 437-440) und Egger (Phytochemistry (1965) 4, 609-618) beschrieben.

45

11

Vorzugsweise sind die Ketocarotinoide ausgewählt aus der Gruppe Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin und Adonixanthin.

5 Ein besonders bevorzugtes Ketocarotinoid ist Astaxanthin.

Die Herstellung der transgenen Pflanzen erfolgt vorzugsweise durch Transformation der Ausgangspflanzen, mit einem Nukleinsäurekonstrukt, das die vorstehend beschriebenen Nukleinsäuren, kodierend eine Ketolase, enthält, die mit einem oder mehreren Regulationssignalen funktionell verknüpft sind, die die Transkription und Translation in Pflanzen gewährleisten.

15 Diese Nukleinsäurekonstrukte, in denen die kodierende Nukleinsäuresequenz mit einem oder mehreren Regulationssignalen funktionell verknüpft sind, die die Transkription und Translation in Pflanzen gewährleisten, werden im folgenden auch Expressionskassetten genannt.

20 Vorzugsweise enthalten die Regulationssignale einen oder mehrere Promotoren, die die Transkription und Translation in Pflanzen gewährleisten.

Die Expressionskassetten beinhalten Regulationssignale, also
25 regulative Nukleinsäuresequenzen, welche die Expression der kodierenden Sequenz in der Wirtszelle steuern. Gemäß einer bevorzugten Ausführungsform umfasst eine Expressionskassette stromaufwärts, d.h. am 5'-Ende der kodierenden Sequenz, einen Promotor und stromabwärts, d.h. am 3'-Ende, ein Polyadenylierungssignal
30 und gegebenenfalls weitere regulatorische Elemente, welche mit der dazwischenliegenden kodierenden Sequenz für mindestens eines der vorstehend beschriebenen Gene operativ verknüpft sind. Unter einer operativen Verknüpfung versteht man die sequenzielle Anordnung von Promotor, kodierender Sequenz, Terminator und ggf. weiterer regulativer Elemente derart, dass jedes der regulativen
35 Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann.

Im folgenden werden beispielhaft die bevorzugten Nukleinsäurekonstrukte, Expressionskassetten und Vektoren für Pflanzen und Verfahren zur Herstellung von transgenen Pflanzen, sowie die transgenen Pflanzen selbst beschrieben.

Die zur operativen Verknüpfung bevorzugten aber nicht darauf
45 beschränkten Sequenzen sind Targeting-Sequenzen zur Gewährleistung der subzellulären Lokalisation im Apoplasten, in der Vakuole, in Plastiden, im Mitochondrium, im Endoplasmatischen Retikulu-

12

lum (ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten und Translationsverstärker wie die 5'-Führungssequenz aus dem Tabak-Mosaik-Virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693-8711).

5

Als Promotoren der Expressionskassette ist grundsätzlich jeder Promotor geeignet, der die Expression von Fremdgenen in Pflanzen steuern kann.

- 10 "Konstitutiver" Promotor meint solche Promotoren, die eine Expression in zahlreichen, bevorzugt allen, Geweben über einen größeren Zeitraum der Pflanzenentwicklung, bevorzugt zu allen Zeitpunkten der Pflanzenentwicklung, gewährleisten.
- 15 Vorzugsweise verwendet man insbesondere einen pflanzlichen Promotor oder einen Promotor, der einem Pflanzenvirus entstammt. Insbesondere bevorzugt ist der Promotor des 35S-Transkriptes des CaMV Blumenkohlmosaikvirus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. (1986) Plant Mol Biol 6:221-228) oder der 19S CaMV Promotor (US 5,352,605; WO 84/02913; Benfey et al. (1989) EMBO J 8:2195-2202).

- Ein weiterer geeigneter konstitutiver Promotor ist der pds Promoter (Pecker et al. (1992) Proc. Natl. Acad. Sci USA 89: 4962-4966) oder der "Rubisco small subunit (SSU)"-Promotor (US 4,962,028), der LeguminB-Promotor (GenBank Acc.-Nr. X03677), der Promotor der Nopalinsynthase aus Agrobacterium, der TR-Doppelpromotor, der OCS (Octopin Synthase) Promotor aus Agrobacterium, der Ubiquitin Promotor (Holtorf S et al. (1995) Plant Mol Biol 29:637-649), den Ubiquitin 1 Promotor (Christensen et al. (1992) Plant Mol Biol 18:675-689; Bruce et al. (1989) Proc Natl Acad Sci USA 86:9692-9696), den Smas Promotor, den Cinnamylalcoholdehydrogenase-Promotor (US 5,683,439), die Promotoren der vakuolärer ATPase Untereinheiten oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991), der Pnit-Promoter (Y07648.L, Hillebrand et al. (1998), Plant. Mol. Biol. 36, 89-99, Hillebrand et al. (1996), Gene, 170, 197-200) sowie weitere Promotoren von Genen, deren konstitutive Expression in Pflanzen dem Fachmann bekannt ist.

- Die Expressionskassetten können auch einen chemisch induzierbaren Promotor enthalten (Übersichtsartikel: Gatz et al. (1997) Annu Rev Plant Physiol Plant Mol Biol 48:89-108), durch den die Expression des Ketolase-Gens in der Pflanze zu einem bestimmten Zeitpunkt gesteuert werden kann. Derartige Promotoren, wie z.B. der PRP1 Promotor (Ward et al. (1993) Plant Mol Biol 22:361-366),

13

durch Salicylsäure induzierbarer Promotor (WO 95/19443), ein durch Benzolsulfonamid-induzierbarer Promotor (EP 0 388 186), ein durch Tetrazyklin-induzierbarer Promotor (Gatz et al. (1992) Plant J 2:397-404), ein durch Abscisinsäure induzierbarer Promotor (EP 0 335 528) bzw. ein durch Ethanol- oder Cyclohexanon-induzierbarer Promotor (WO 93/21334) können ebenfalls verwendet werden.

Ferner sind Promotoren bevorzugt, die durch biotischen oder abiotischen Stress induziert werden wie beispielsweise der pathogen-induzierbare Promotor des PRP1-Gens (Ward et al. (1993) Plant Mol Biol 22:361-366), der hitzeinduzierbare hsp70- oder hsp80-Promoter aus Tomate (US 5,187,267), der kälteinduzierbare alpha-Amylase Promoter aus der Kartoffel (WO 96/12814), der licht-induzierbare PPKK Promotor oder der verwundungsinduzierte pinII-Promoter (EP375091).

Pathogen-induzierbare Promotoren umfassen die von Genen, die infolge eines Pathogenbefalls induziert werden wie beispielsweise Gene von PR-Proteinen, SAR-Proteinen, b-1,3-Glucanase, Chitinase usw. (beispielsweise Redolfi et al. (1983) Neth J Plant Pathol 89:245-254; Uknes, et al. (1992) The Plant Cell 4:645-656; Van Loon (1985) Plant Mol Biol 4:111-116; Marieneau et al. (1987) Plant Mol Biol 9:335-342; Matton et al. (1987) Molecular Plant-Microbe Interactions 2:325-342; Somssich et al. (1986) Proc Natl Acad Sci USA 83:2427-2430; Somssich et al. (1988) Mol Gen Genetics 2:93-98; Chen et al. (1996) Plant J 10:955-966; Zhang and Sing (1994) Proc Natl Acad Sci USA 91:2507-2511; Warner, et al. (1993) Plant J 3:191-201; Siebertz et al. (1989) Plant Cell 1:961-968(1989).

Umfasst sind auch verwundungs-induzierbare Promotoren wie der des pinII Gens (Ryan (1990) Ann Rev Phytopath 28:425-449; Duan et al. (1996) Nat Biotech 14:494-498), des wun1 und wun2-Gens (US 5,428,148), des win1- und win2-Gens (Stanford et al. (1989) Mol Gen Genet 215:200-208), des Systemin (McGurl et al. (1992) Science 225:1570-1573), des WIP1-Gens (Rohmeier et al. (1993) Plant Mol Biol 22:783-792; Ekelkamp et al. (1993) FEBS Letters 323:73-76), des MPI-Gens (Corderok et al. (1994) The Plant J 6(2):141-150) und dergleichen.

Weitere geeignete Promotoren sind beispielsweise fruchtreifungsspezifische Promotoren, wie beispielsweise der fruchtreifungsspezifische Promotor aus Tomate (WO 94/21794, EP 409 625). Entwicklungsabhängige Promotoren schließen zum Teil die gewebespezifischen Promotoren ein, da die Ausbildung einzelner Gewebe naturgemäß entwicklungsabhängig erfolgt.

14

Weiterhin sind insbesondere solche Promotoren bevorzugt, die die Expression in Geweben oder Pflanzenteilen sicherstellen, in denen beispielsweise die Biosynthese von Ketocarotinoiden bzw. dessen Vorstufen stattfindet. Bevorzugt sind beispielsweise Promotoren mit Spezifitäten für die Antheren, Ovarien, Petalen, Sepalen, Blüten, Blätter, Stengel, Wurzeln und Früchte und Kombinationen hieraus.

Knollen-, Speicherwurzel- oder Wurzel-spezifische Promotoren sind beispielsweise der Patatin Promotor Klasse I (B33) oder der Promotor des Cathepsin D Inhibitors aus Kartoffel.

Blattspezifische Promotoren sind beispielsweise der Promotor der cytosolischen FBPase aus Kartoffel (WO 97/05900), der SSU Promotor (small subunit) der Rubisco (Ribulose-1,5-bisphosphat-carboxylase) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al. (1989) EMBO J 8:2445-2451).

Blütenspezifische Promotoren sind beispielsweise der Phytoen Synthase Promotor (WO 92/16635) oder der Promotor des P-rr Gens (WO 98/22593).

Antheren-spezifische Promotoren sind beispielsweise der 5126-Promotor (US 5,689,049, US 5,689,051), der glob-1 Promotor oder der g-Zein Promotor.

Fruchtspezifische Promotoren sind beispielsweise

der Pds-Promoter aus Tomate (Genbank-ACCESSION U46919; Corona, V., Aracri, B., Kosturkova, G., Bartley, G.E., Pitto, L., Giorgetti, L., Scolnik, P.A. and Giuliano, G., Regulation of a carotenoid biosynthesis gene promoter during plant development Plant J. 9 (4), 505-512 (1996)), SEQ ID NO.17,

der 2A11 Promoter aus Tomate (Pear, J.R., Ridge, N., Rasmussen, R., Rose, R.E. and Houck, C.M. Isolation and characterization of a fruit-specific cDNA and the corresponding genomic clone from tomato Plant Mol. Biol. 13 (6), 639-651 (1989), SEQ ID NO. 18,

der Cucumisin Promoter (Yamagata, H., Yonesu, K., Hirata, A. and Aizono, Y., TGTCACA Motif Is a Novel cis-Regulatory Enhancer Element Involved in Fruit-specific Expression of the cucumisin Gene J. Biol. Chem. 277 (13), 11582-11590 (2002), SEQ ID NO. 19,

15

der Promoter des Endogalacturonasegens (Redondo-Nevado, J., Medina-Escobar, N., Caballero-Repullo, J.L. and Munoz-Blanco, J.

A fruit-specific and developmentally regulated endo-polygalacturonase gene from strawberry (*Fragaria x ananassa* c.v. Chandler), J Experimental Botany 52 (362) 1941-1945 (2001), SEQ ID NO. 20,

der Polygalacturonase Promoter aus Tomate (Nicholass, F.J., Smith, C.J., Schuch, W., Bird, C.R. and Grierson, D., High levels of ripening-specific reporter gene expression directed by tomato fruit polygalacturonase gene-flanking regions, Plant Mol. Biol. 28 (3), 423-435 (1995)), SEQ ID NO. 21,

die TMF7 und TMF9 Promotoren (US 5608150),

der Promotor E4 (Cordes S. Deikman J. Margossian LJ. Fischer RL. Interaction of a developmentally regulated DNA-binding factor with sites flanking two different fruit-ripening genes from tomato (1989), Plant Cell 1, 1025-1034) und

der Promotor E8 (Deikman and Fisher, Interaction of a DNA binding factor with the 5'-flanking region of an ethylene-responsive fruit ripening gene from tomato (1988), EMBO J. 7, 3315-3320). Weitere zur Expression in Pflanzen geeignete Promotoren sind beschrieben (Rogers et al. (1987) Meth in Enzymol 153:253-277; Schardl et al. (1987) Gene 61:1-11; Berger et al. (1989) Proc Natl Acad Sci USA 86:8402-8406).

Alle in der vorliegenden Anmeldung beschriebenen Promotoren ermöglichen in der Regel die Expression der Ketolase in Früchten der erfindungsgemäßen Pflanzen.

Besonders bevorzugt im erfindungsgemäßen Verfahren sind konstitutive sowie insbesondere fruchtspezifische Promotoren.

Die vorliegende Erfindung betrifft daher insbesondere ein Nukleinsäurekonstrukt, enthaltend funktionell verknüpft einen fruchtspezifischen Promotor, besonders bevorzugt einen oben beschriebenen fruchtspezifischen Promotor, und eine Nukleinsäure, kodierend eine Ketolase.

Die Herstellung einer Expressionskassette erfolgt vorzugsweise durch Fusion eines geeigneten Promotors mit einer vorstehend beschriebenen Nukleinsäure kodierend eine Ketolase und vorzugsweise einer zwischen Promotor und Nukleinsäure-Sequenz inserierten Nukleinsäure, die für ein plastidenspezifisches Transitpeptid kodiert, sowie einem Polyadenylierungssignal nach gängigen

16

Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1987) beschrieben sind.

- 10 Die vorzugsweise insertierte Nukleinsäuren, kodierend ein plastidäres Transitpeptid, gewährleisten die Lokalisation in Plastiden und insbesondere in Chromoplasten.

Es können auch Expressionskassetten verwendet werden, deren Nukleinsäure-Sequenz für ein Ketolase-Fusionsprotein kodiert, wobei ein Teil des Fusionsproteins ein Transitpeptid ist, das die Translokation des Polypeptides steuert. Bevorzugt sind für die Chromoplasten spezifische Transitpeptide, welche nach Translokation der Ketolase in die Chromoplasten vom Ketolase-Teil enzymatisch abgespalten werden.

Insbesondere bevorzugt ist das Transitpeptid, das von der plastidären *Nicotiana tabacum* Transketolase oder einem anderen Transitpeptid (z.B. dem Transitpeptid der kleinen Untereinheit der Rubisco (rbcS) oder der Ferredoxin NADP Oxidoreduktase als auch der Isopentenylpyrophosphat Isomerase-2) oder dessen funktionellem Äquivalent abgeleitet ist.

Besonders bevorzugt sind Nukleinsäure-Sequenzen von drei Kassetten des Plastiden-Transitpeptids der plastidären Transketolase aus Tabak in drei Leserastern als KpnI/BamHI Fragmente mit einem ATG-Kodon in der NcoI Schnittstelle:

pTP09

35

KpnI_GGTACCATGGCGTCTTCTTCTTCTCTCACTCTCTCTCAAGCTATCCTCTCTCGTTCTGTC
CCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCTCACTTTTTCCGGCCTTAA
ATCCAATCCCAATATCACCACTCCCGCCGCGTACTCCTTCCTCCGCCGCCGCCGCCGCCGTCG
TAAGGTCACCGGCGATTTCGTGCCTCAGCTGCAACCGAAACCATAGAGAAAAGTGAAGTGCAGGGA

40 TCC_BamHI

pTP10

KpnI_GGTACCATGGCGTCTTCTTCTTCTCTCACTCTCTCTCAAGCTATCCTCTCTCGTTCTGTC
45 CCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCTCACTTTTTCCGGCCTTAA
ATCCAATCCCAATATCACCACTCCCGCCGCGTACTCCTTCCTCCGCCGCCGCCGCCGCCGTCG

17

TAAGGTCACCGGCGATTTCGTGCCTCAGCTGCAACCGAAACCATAGAGAAAACCTGAGACTGCGCTG
GATCC_BamHI

pTP11

5

KpnI_GGTACCATGGCGTCTTCTTCTTCTCTCAAGCTATCCTCTCTCGTTCTGTC
CCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCTTCTTCTCTCACTTTTCCGGCCTTAA
ATCCAATCCCAATATCACCACCTCCCGCCGCGTACTCCTTCCTCCGCGCCGCGCCGCGCTCG
TAAGGTCACCGGCGATTTCGTGCCTCAGCTGCAACCGAAACCATAGAGAAAACCTGAGACTGCGGGG

10 ATCC_BamHI

Weitere Beispiele für ein plastidäres Transitpeptid sind das Transitpeptid der plastidären Isopentenyl-pyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* und das Transitpeptid der kleinen Untereinheit der Ribulosebisdiphosphat Carboxylase (rbcS) aus Erbse (Guerineau, F, Woolston, S, Brooks, L, Mullineaux, P (1988) An expression cassette for targeting foreign proteins into the chloroplasts. Nucl. Acids Res. 16: 11380).

20 Die erfindungsgemäßen Nukleinsäuren können synthetisch hergestellt oder natürlich gewonnen sein oder eine Mischung aus synthetischen und natürlichen Nukleinsäure-Bestandteilen enthalten, sowie aus verschiedenen heterologen Genabschnitten verschiedener Organismen bestehen.

25

Bevorzugt sind, wie vorstehend beschrieben, synthetische Nukleotid-Sequenzen mit Kodons, die von Pflanzen bevorzugt werden. Diese von Pflanzen bevorzugten Kodons können aus Kodons mit der höchsten Proteinhäufigkeit bestimmt werden, die in den meisten

30 interessanten Pflanzenspezies exprimiert werden.

Bei der Präparation einer Expressionskassette können verschiedene DNA-Fragmente manipuliert werden, um eine Nukleotid-Sequenz zu erhalten, die zweckmäßigerweise in der korrekten Richtung liest und die mit einem korrekten Leseraster ausgestattet ist. Für die Verbindung der DNA-Fragmente miteinander können an die Fragmente

35 Adaptoren oder Linker angesetzt werden.

Zweckmäßigerweise können die Promotor- und die Terminator-Regionen in Transkriptionsrichtung mit einem Linker oder Polylinker, der eine oder mehrere Restriktionsstellen für die Insertion dieser Sequenz enthält, versehen werden. In der Regel hat der Linker 1 bis 10, meistens 1 bis 8, vorzugsweise 2 bis 6 Restriktionsstellen. Im allgemeinen hat der Linker innerhalb der regulatorischen Bereiche eine Größe von weniger als 100 bp, häufig weniger als 60 bp, mindestens jedoch 5 bp. Der Promotor kann sowohl nativ bzw. homolog als auch fremdartig bzw. heterolog zur Wirtspflanze

40

45

18

sein. Die Expressionskassette beinhaltet vorzugsweise in der 5'-3'-Transkriptionsrichtung den Promotor, eine kodierende Nukleinsäuresequenz oder ein Nukleinsäurekonstrukt und eine Region für die transkriptionale Termination. Verschiedene Terminationsbereiche sind gegeneinander beliebig austauschbar.

Ein Beispiel für einen Terminator ist der 35S-Terminator (Guerineau et al. (1988) Nucl Acids Res. 16: 11380), der nos Terminator (Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman
10 HM. Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet. 1982;1(6):561-73) oder der ocs Terminator (Gielen, J, de Beuckeleer, M, Seurinck, J, Debroek, H, de Greve, H, Lemmers, M, van Montagu, M, Schell, J (1984) The complete sequence of the TL-DNA of the Agrobacterium tumefaciens plasmid
15 pTiAch5. EMBO J. 3: 835-846).

Ferner können Manipulationen, die passende Restriktionsschnittstellen bereitstellen oder die überflüssige DNA oder Restriktionsschnittstellen entfernen, eingesetzt werden. Wo Insertionen,
20 Deletionen oder Substitutionen wie z.B. Transitionen und Transversionen in Frage kommen, können *in vitro*-Mutagenese, "primer-repair", Restriktion oder Ligation verwendet werden.

Bei geeigneten Manipulationen, wie z.B. Restriktion, "chewing-back" oder Auffüllen von Überhängen für "bluntends", können
25 komplementäre Enden der Fragmente für die Ligation zur Verfügung gestellt werden.

Bevorzugte Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA-Polyadenylierungssignale aus Agrobacterium tumefaciens, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACH5 entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff)
30 oder funktionelle Äquivalente.

35 Die Übertragung von Fremdgenen in das Genom einer Pflanze wird als Transformation bezeichnet.

Dazu können an sich bekannte Methoden zur Transformation und
40 Regeneration von Pflanzen aus Pflanzengeweben oder Pflanzenzellen zur transienten oder stabilen Transformation genutzt werden.

Geeignete Methoden zur Transformation von Pflanzen sind die Protoplastentransformation durch Polyethylenglykol-induzierte DNA-
45 Aufnahme, das biolistische Verfahren mit der Genkanone - die sogenannte particle bombardment Methode, die Elektroporation, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Mikro-

19

injektion und der, vorstehend beschriebene, durch Agrobacterium vermittelte Gentransfer. Die genannten Verfahren sind beispielsweise in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press (1993), 128-143 sowie in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225) beschrieben.

Vorzugsweise wird das zu exprimierende Konstrukt in einen Vektor kloniert, der geeignet ist, Agrobacterium tumefaciens zu transformieren, beispielsweise pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711) oder besonders bevorzugt pSUN2, pSUN3, pSUN4 oder pSUN5 (WO 02/00900).

Mit einem Expressionsplasmid transformierte Agrobakterien können in bekannter Weise zur Transformation von Pflanzen verwendet werden, z.B. indem verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

20

Zur bevorzugten Herstellung von genetisch veränderten Pflanzen, im folgenden auch transgene Pflanzen bezeichnet, wird die fusionierte Expressionskassette, die eine Ketolase exprimiert, in einen Vektor, beispielsweise pBin19 oder insbesondere pSUN2 kloniert, der geeignet ist, in Agrobacterium tumefaciens transformiert zu werden

Mit einem solchen Vektor transformierte Agrobakterien können dann in bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen verwendet werden, indem beispielsweise verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

Die Transformation von Pflanzen durch Agrobakterien ist unter anderem bekannt aus F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, 1993, S. 15-38. Aus den transformierten Zellen der verwundeten Blätter bzw. Blattstücke können in bekannter Weise transgene Pflanzen regeneriert werden, die ein in die Expressionskassette integriertes Gen für die Expression einer Nukleinsäure codierend eine Ketolase enthalten.

Zur Transformation einer Wirtspflanze mit einer für eine Ketolase kodierenden Nukleinsäure wird eine Expressionskassette als Insertion in einen rekombinanten Vektor eingebaut, dessen Vektor-DNA zusätzliche funktionelle Regulationssignale, beispielsweise

20

Sequenzen für Replikation oder Integration enthält. Geeignete Vektoren sind unter anderem in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Kap. 6/7, S. 71-119 (1993) beschrieben.

5

Unter Verwendung der oben zitierten Rekombinations- und Klonierungstechniken können die Expressionskassetten in geeignete Vektoren kloniert werden, die ihre Vermehrung, beispielsweise in *E. coli*, ermöglichen. Geeignete Klonierungsvektoren sind u.a.

- 10 pJIT117 (Guerineau et al. (1988) Nucl. Acids Res. 16 :11380), pBR332, pUC-Serien, M13mp-Serien, pACYC184, pMC1210, pMcl 210 und pCL1920. Besonders geeignet sind binäre Vektoren, die sowohl in *E. coli* als auch in Agrobakterien replizieren können.

- 15 Dabei kann je nach Wahl des Promotors die Expression konstitutiv oder vorzugsweise spezifisch in den Früchten erfolgen.

- Dementsprechend betrifft die Erfindung ferner ein Verfahren zur Herstellung von genetisch veränderten Pflanzen, dadurch gekennzeichnet, dass man ein Nukleinsäurekonstrukt, enthaltend funktionell verknüpft einen fruchtspezifischen Promotor und Nukleinsäuren kodierend eine Ketolase in das Genom der Ausgangspflanze einführt.
- 20

- 25 Die Erfindung betrifft ferner die genetisch veränderten Pflanzen, die im Vergleich zur Ausgangspflanze in Früchten eine Ketolase-Aktivität aufweist.

- Die Ketolaseaktivität wird in einer bevorzugten Ausführungsform dadurch erreicht, dass die genetisch veränderte Pflanze in den Früchten eine Ketolase exprimiert.
- 30

- Die bevorzugten, genetisch veränderten Pflanzen enthalten daher in Früchten mindestens eine Nukleinsäure, kodierend eine Ketolase.
- 35

- In einer weiter bevorzugten Ausführungsform erfolgt, wie vorstehend ausgeführt, die Verursachung der Genexpression einer Nukleinsäure, kodierend eine Ketolase, durch Einbringen von Nukleinsäuren, kodierend eine Ketolase, in die Ausgangspflanze.
- 40

- Der Erfindung betrifft daher besonders bevorzugt eine vorstehend beschriebene genetisch veränderte Pflanze, dadurch gekennzeichnet, dass man in die Pflanze ausgehend von einer Ausgangspflanze mindestens eine Nukleinsäure, kodierend eine Ketolase eingebracht hat.
- 45

21

Die Erfindung betrifft insbesondere genetisch veränderte Pflanzen, ausgewählt aus den Pflanzengattungen Actinophloeus, Aglaeonema, Ananas, Arbutus, Archontophoenix, Area, Aronia, Asparagus, Attalea, Berberis, Bixia, Brachychilum, Bryonia, Caliptocalix, 5 Capsicum, Carica, Celastrus, Citrullus, Citrus, Convallaria, Cotoneaster, Crataegus, Cucumis, Cucurbita, Cuscuta, Cycas, Cyphomandra, Dioscorea, Diospyrus, Dura, Elaeagnus, Elaeis, Erythroxylon, Euonymus, Ficus, Fortunella, Fragaria, Gordinia, Gonocaryum, Gossypium, Guava, Guilielma, Hibiscus, Hippophaea, 10 Iris, Lathyrus, Lonicera, Luffa, Lycium, Lycopersicum, Malpighia, Mangifera, Mormodica, Murraya, Musa, Nenga, Palisota, Pandanus, Passiflora, Persea, Physalis, Prunus, Ptychandra, Punica, Pyracantha, Pyrus, Ribes, Rosa, Rubus, Sabal, Sambucus, Seaforita, Shepherdia, Solanum, Sorbus, Synaspadix, Tabernae, Tamus, Taxus, 15 Trichosanthes, Triphasia, Vaccinium, Viburnum, Vignia oder Vitis, enthaltend mindestens eine Nukleinsäure, kodierend eine Ketolase.

Ganz besonders bevorzugte Pflanzengattungen sind Ananas, Asparagus, Capsicum, Citrus, Cucumis, Cucurbita, Citrullus, 20 Lycopersicum, Passiflora, Prunus, Physalis, Solanum, Vaccinium und Vitis, enthaltend mindestens eine transgene Nukleinsäure, kodierend eine Ketolase.

Wie vorstehend erwähnt wird in bevorzugten transgenen Pflanzen 25 die Ketolase in den Früchten exprimiert, besonderes bevorzugt ist die Expression der Ketolase in den Früchten am höchsten.

Die transgenen Pflanzen, deren Vermehrungsgut, sowie deren Pflanzenzellen, -gewebe oder -teile, insbesondere deren Früchte sind 30 ein weiterer Gegenstand der vorliegenden Erfindung.

Die genetisch veränderten Pflanzen können, wie vorstehend beschrieben, zur Herstellung von Ketocarotinoiden, insbesondere Astaxanthin, verwendet werden.

35

Von Menschen und Tieren verzehrbare erfindungsgemäße, genetisch veränderte Pflanzen mit erhöhtem Gehalt an Ketocarotinoiden können auch beispielsweise direkt oder nach an sich bekannter Prozessierung als Nahrungsmittel oder Futtermittel oder als Futter- 40 und Nahrungsergänzungsmittel verwendet werden. Ferner können die genetisch veränderten Pflanzen zur Herstellung von Ketocarotinoid-haltigen Extrakten der Pflanzen und/oder zur Herstellung von Futter- und Nahrungsergänzungsmitteln verwendet werden.

45 Die genetisch veränderten Pflanzen weisen im Vergleich zum Wildtyp einen erhöhten Gehalt an Ketocarotinoiden auf.

22

Unter einem erhöhten Gehalt an Ketocarotinoiden wird in der Regel ein erhöhter Gehalt an Gesamt-Ketocarotinoid verstanden.

Unter einem erhöhten Gehalt an Ketocarotinoiden wird aber auch
5 insbesondere ein veränderter Gehalt der bevorzugten Ketocarotinoide verstanden, ohne dass zwangsläufig der Gesamt-Carotinoidgehalt erhöht sein muss.

In einer besonders bevorzugten Ausführungsform weisen die
10 erfindungsgemäßen, genetisch veränderten Pflanzen im Vergleich zum Wildtyp einen erhöhten Gehalt an Astaxanthin auf.

Unter einem erhöhten Gehalt wird in diesem Fall insbesondere ein verursachter Gehalt an Ketocarotinoiden, bzw. Astaxanthin
15 verstanden.

Die Erfindung wird durch die nun folgenden Beispiele erläutert, ist aber nicht auf diese beschränkt:

20 Allgemeine Experimentelle Bedingungen:
Sequenzanalyse rekombinanter DNA

Die Sequenzierung rekombinanter DNA-Moleküle erfolgte mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma Licor (Vertrieb durch
25 MWG Biotech, Ebersbach) nach der Methode von Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

Beispiel 1: Amplifikation einer cDNA, die die gesamte Primärsequenz der Ketolase aus *Haematococcus pluvialis* Flo-
30 tow em. Wille kodiert

Die cDNA, die für die Ketolase aus *Haematococcus pluvialis* kodiert, wurde mittels PCR aus *Haematococcus pluvialis* (Stamm 192.80 der "Sammlung von Algenkulturen der Universität Göttingen") Suspensionskultur amplifiziert.
35

Für die Präparation von Total-RNA aus einer Suspensionskultur von *Haematococcus pluvialis* (Stamm 192.80), die 2 Wochen mit indirektem Tageslicht bei Raumtemperatur in *Haematococcus*-Medium
40 (1.2 g/l Natriumacetat, 2 g/l Hefeextrakt, 0.2 g/l $\text{MgCl}_2 \times 6\text{H}_2\text{O}$, 0.02 $\text{CaCl}_2 \times 2\text{H}_2\text{O}$; pH 6.8; nach Autoklavieren Zugabe von 400 mg/l L-Asparagin, 10 mg/l $\text{FeSO}_4 \times \text{H}_2\text{O}$) gewachsen war, wurden die Zellen geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert. Anschließend wurden 100 mg der gefrorenen, pulverisierten Algenzellen in ein Reaktionsgefäß überführt und in 0,8 ml Trizol-Puffer (Life Technologies) aufgenommen. Die Suspension
45 wurde mit 0,2 ml Chloroform extrahiert. Nach 15minütiger Zentri-

23

fugation bei 12000 g wurde der wässrige Überstand abgenommen und in ein neues Reaktionsgefäß überführt und mit einem Volumen Ethanol extrahiert. Die RNA wurde mit einem Volumen Isopropanol gefällt, mit 75 % Ethanol gewaschen und das Pellet in DEPC Wasser
5 (über Nacht Inkubation von Wasser mit 1/1000 Volumen Diethylpyrocarbonat bei Raumtemperatur, anschließend autoklaviert) gelöst. Die RNA-Konzentration wurde photometrisch bestimmt.

Für die cDNA-Synthese wurden 2.5 µg Gesamt-RNA für 10 min bei
10 60°C denaturiert, für 2 min auf Eis abgekühlt und mittels eines cDNA-Kits (Ready-to-go-you-prime-beads, Pharmacia Biotech) nach Herstellerangaben unter Verwendung eines antisense spezifischen Primers (PR1 SEQ ID No. 29) in cDNA umgeschrieben.

15 Die Nukleinsäure codierend eine Ketolase aus *Haematococcus pluvialis* (Stamm 192.80) wurde mittels polymerase chain reaction (PCR) aus *Haematococcus pluvialis* unter Verwendung eines sense spezifischen Primers (PR2 SEQ ID No. 30) und eines antisense spezifischen Primers (PR1 SEQ ID No. 29) amplifiziert.

20

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der cDNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in
25 einem 50 µl Reaktionsansatz, in dem enthalten war:

- 4 µl einer *Haematococcus pluvialis* cDNA (hergestellt wie oben beschrieben)
- 0,25 mM dNTPs
- 30 - 0,2 mM PR1 (SEQ ID No. 29)
- 0,2 mM PR2 (SEQ ID No. 30)
- 5 µl 10X PCR-Puffer (TAKARA)
- 0,25 µl R Taq Polymerase (TAKARA)
- 25,8 µl Aq. Dest.

35

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

	1X	94°C	2 Minuten
	35X	94°C	1 Minute
40		53°C	2 Minuten
		72°C	3 Minuten
	1X	72°C	10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 29 und SEQ ID No. 30 resul-
45 tierte in einem 1155 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (SEQ ID No. 22). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-

24

Klonierungsvektor pGEM-Teasy (Promega) kloniert und der Klon pGKETO2 erhalten.

Sequenzierung des Klons pGKETO2 mit dem T7- und dem SP6-Primer
5 bestätigte eine Sequenz, die sich lediglich in den drei Ko-
dons 73, 114 und 119 in je einer Base von der publizierten Se-
quenz X86782 unterscheidet. Diese Nukleotidaustausche wurden
in einem unabhängigen Amplifikationsexperiment reproduziert
und repräsentieren somit die Nukleotidsequenz im verwendeten *Haem-*
10 *matococcus pluvialis* Stamm 192.80 (Abbildung 3 und 4, Sequenzver-
gleiche).

Dieser Klon wurde daher für die Klonierung in den Expressionsvek-
tor pJIT117 (Guerineau et al. 1988, Nucl. Acids Res. 16: 11380)
15 verwendet. Die Klonierung erfolgte durch Isolierung des 1027 Bp
SpHI-Fragmentes aus pGKETO2 und Ligierung in den SpHI geschnitten-
en Vektor pJIT117. Der Klon, der das *Haematococcus pluvialis* Ke-
tolasegen in der korrekten Orientierung als N-terminale transla-
tionale Fusion mit der rbcS Transitpeptidsequenz enthält, heißt
20 pJKETO2.

Beispiel 2: Amplifikation einer cDNA, die die Ketolase aus *Haema-*
tococcus pluvialis Flotow em. Wille mit einem um 14
Aminosäuren verkürztem N-terminus kodiert

25

Die cDNA, die für die Ketolase aus *Haematococcus pluvialis*
(Stamm 192.80) mit einem um 14 Aminosäuren verkürztem N-Termi-
nus kodiert, wurde mittels PCR aus *Haematococcus pluvialis* Sus-
pensionskultur (Stamm 192.80 der "Sammlung von Algenkulturen der
30 Universität Göttingen") amplifiziert.

Die Präparation von Total-RNA aus einer Suspensionskultur von
Haematococcus pluvialis (Stamm 192.80) erfolgte wie in Beispiel 1
beschrieben.

35 Die cDNA-Synthese erfolgte wie unter Beispiel 1 beschrieben.

Die Nukleinsäure, kodierend eine Ketolase, aus *Haematococcus*
pluvialis (Stamm 192.80) mit einem um 14 Aminosäuren verkürztem
N-Terminus wurde mittels polymerase chain reaction (PCR) aus
40 *Haematococcus pluvialis* unter Verwendung eines sense spezifischen
Primers (PR3 SEQ ID No. 31) und eines antisense spezifischen Pri-
mers (PR1 SEQ ID No. 29) amplifiziert.

45

25

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der cDNA, die für ein Ketolase Protein mit um 14 Aminosäuren verkürztem N-Terminus kodiert, erfolgte in
5 einem 50 µl Reaktionsansatz, in dem enthalten war:

- 4 µl einer *Haematococcus pluvialis* cDNA (hergestellt wie oben beschrieben)
- 0,25 mM dNTPs
- 10 - 0,2 mM PR1 (SEQ ID No. 29)
- 0,2 mM PR3 (SEQ ID No. 31)
- 5 µl 10X PCR-Puffer (TAKARA)
- 0,25 µl R Taq Polymerase (TAKARA)
- 25,8 µl Aq. Dest.

15

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

	1X	94°C	2 Minuten
	35X	94°C	1 Minute
20		53°C	2 Minuten
		72°C	3 Minuten
	1X	72°C	10 Minuten

Die PCR-Amplifikation mit SEQ ID No.29 und SEQ ID No. 31 resultierte in einem 1111 Bp Fragment, das für ein Ketolase Protein kodiert, bei dem N-terminalen Aminosäuren (Position 2-16) durch eine einzige Aminosäure (Leucin) ersetzt sind.
25

Das Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pGEM-Teasy (Promega) kloniert und der Klon pGKETO3 erhalten. Sequenzierungen mit den Primern T7- und SP6 bestätigten eine zur Sequenz SEQ ID No. 22 identische Sequenz, wobei die 5'Region (Position 1-53) der SEQ ID No. 22 im Amplifikat SEQ ID No. 24 durch eine in der Sequenz abweichende Nonamersequenz ersetzt wurde. Dieser Klon wurde daher für die Klonierung in den Expressionsvektor pJIT117 (Guerineau et al. 1988, Nucl. Acids Res. 16: 11380) verwendet.
30
35

Die Klonierung erfolgte durch Isolierung des 985 Bp SpHI Fragmentes aus pGKETO3 und Ligierung mit dem SpHI geschnittenen Vektor pJIT117. Der Klon, der die *Haematococcus pluvialis* Ketolase mit einem um 14 Aminosäuren verkürztem N-Terminus in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heißt pJKETO3.
40
45

26

Beispiel 3: Amplifikation einer cDNA, die die Ketolase aus *Haematococcus pluvialis* Flotow em. Wille (Stamm 192.80 der "Sammlung von Algenkulturen der Universität Göttingen") bestehend aus der gesamten Primärsequenz und fusioniertem C-terminalem myc-Tag kodiert

5

Die cDNA, die für die Ketolase aus *Haematococcus pluvialis* (Stamm 192.80) bestehend aus der gesamten Primärsequenz und fusioniertem C-terminalem myc-Tag kodiert, wurde mittels PCR unter Verwendung des Plasmids pGKETO2 (in Beispiel 1 beschrieben) und des Primers PR15 (SEQ ID No. 32) hergestellt. Der Primer PR15 setzt sich zusammen aus einer antisense spezifischen 3'Region (Nucleotide 40-59) und einer myc-Tag kodierenden 5'Region (Nucleotide 1-39).

15 Die Denaturierung (5 min bei 95°C) und Annealing (langsame Abkühlung bei Raumtemperatur auf 40°C) von pGKETO2 und PR15 erfolgte in einem 11,5 µl Reaktionsansatz, in dem enthalten war:

- 1 µg pGKETO2 PlasmidDNA
- 20 - 0,1 µg PR15 (SEQ ID No. 32)

Das Auffüllen der 3'Enden (30 min bei 30°C) erfolgte in einem 20 µl Reaktionsansatz, in dem enthalten war:

- 25 - 11,5 µl pGKETO2/PR15-Annealingsreaktion (hergestellt wie oben beschrieben)
- 50 µM dNTPs
- 2 µl 1X Klenow Puffer
- 2U Klenow Enzym

30

Die Nukleinsäure kodierend eine Ketolase aus *Haematococcus pluvialis* (Stamm 192.80) bestehend aus der gesamten Primärsequenz und fusioniertem C-terminalem myc-Tag wurde mittels polymerase chain reaction (PCR) aus *Haematococcus pluvialis* unter Verwendung eines sense spezifischen Primers (PR2 SEQ ID No. 30) und eines antisense spezifischen Primers (PR15 SEQ ID No. 32) amplifiziert.

35

Die PCR-Bedingungen waren die folgenden:

40 Die PCR zur Amplifikation der cDNA, die für ein Ketolase Protein mit fusioniertem C-terminalem myc-Tag kodiert, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 1 µl einer Annealingsreaktion (hergestellt wie oben beschrieben)
- 45 - 0,25 mM dNTPs
- 0,2 µM PR15 (SEQ ID No. 32)

27

- 0,2 µM PR2 (SEQ ID No. 30)
- 5 µl 10X PCR-Puffer (TAKARA)
- 0,25 µl R Taq Polymerase (TAKARA)
- 28,8 µl Aq. Dest.

5

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

	1X	94°C	2 Minuten
	35X	94°C	1 Minute
10		53°C	1 Minute
		72°C	1 Minute
	1X	72°C	10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 32 und SEQ ID No. 30 resultierte in einem 1032 Bp-Fragment, das für ein Protein kodiert, bestehend aus der gesamten Primärsequenz der Ketolase aus *Haematococcus pluvialis* als zweifache translationale Fusion mit dem rbcS Transitpeptide am N-Terminus und dem myc-Tag am C-Terminus.

20 Das Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pGEM-Teasy (Promega) kloniert, und der Klon pGKET04 erhalten. Sequenzierungen mit den Primern T7- und SP6 bestätigten eine zur Sequenz SEQ ID No. 22 identische Sequenz, wobei die 3'Region (Position 993-1155) der SEQ ID No. 22 im Amplifikat SEQ ID No. 26 durch eine in der abweichende Sequenz aus 39 Bp ersetzt wurde. Dieser Klon wurde daher für die Klonierung in den Expressionsvektor pJIT117 (Guerineau et al. 1988, Nucl. Acids Res. 16: 11380) verwendet.

30 Die Klonierung erfolgte durch Isolierung des 1038 Bp EcoRI-SpHI Fragmentes aus pGKET04 und Ligierung mit dem EcoRI-SpHI geschnittenen Vektor pJIT117. Durch die Ligation entsteht eine translationale Fusion zwischen dem C-Terminus der rbcS Transitpeptidsequenz und dem N-Terminus der Ketolase Sequenz. Der Klon, der 35 die *Haematococcus pluvialis* Ketolase mit fusioniertem C-terminalem myc-Tag in der korrekten Orientierung als translationale N-terminale Fusion mit dem rbcS Transitpeptid enthält, heißt pJKET4.

40 Beispiel 4: Herstellung von Expressionsvektoren zur konstitutiven Expression der *Haematococcus pluvialis* Ketolase in *Lycopersicon esculentum*

Die Expression der Ketolase aus *Haematococcus pluvialis* in 45 *L. esculentum* erfolgte unter Kontrolle des konstitutiven Promoters d35S aus CaMV (Franck et al. 1980, Cell 21: 285-294). Die

28

Expression erfolgte mit dem Transitpeptid *rbcS* aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715).

Die Herstellung eines Expressionsplasmides für die Agrobacterium-vermittelte Transformation der Ketolase aus *Haematococcus pluvialis* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

10 - Zur Herstellung des Expressionsvektors pS3KETO2 wurde das 2.8 Kb SacI-XhoI Fragment aus pJKETO2 mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert (Abbildung 5, Konstruktkarte). In der Abbildung 5 beinhaltet Fragment d35S den duplizierten 35S Promoter (747 bp), Fragment *rbcS* das *rbcS* Transitpeptid aus Erbse (204 bp), Fragment KETO2 (1027 bp) die gesamte Primärsequenz kodierend für die *Haematococcus pluvialis* Ketolase, 15 Fragment *term* (761 bp) das Polyadenylierungssignal von CaMV.

20 - Zur Herstellung des Expressionsvektors pS3KETO3 wurde das 2.7 Kb bp SacI-XhoI Fragment aus pJKETO3 mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert. (Abbildung 6, Konstruktkarte). In der Abbildung 6 beinhaltet Fragment d35S den duplizierten 35S Promoter (747 bp), Fragment *rbcS* das *rbcS* Transitpeptid aus Erbse (204 bp), Fragment KETO3 (985 bp) die um 14 N-terminale Aminosäuren verkürzte Primärsequenz kodierend für die 25 *Haematococcus pluvialis* Ketolase, Fragment *term* (761 bp) das Polyadenylierungssignal von CaMV.

Zur Herstellung des Expressionsvektors pS3KETO4 wurde das 2.8 Kb SacI-XhoI Fragment aus pJKETO4 mit dem SacI-XhoI geschnittenen 30 Vektor pSUN3 ligiert. (Abbildung 7, Konstruktkarte). In der Abbildung 7 beinhaltet Fragment d35S den duplizierten 35S Promoter ((747 bp), Fragment *rbcS* das *rbcS* Transitpeptid aus Erbse (204 bp), Fragment KETO4 (1038 bp) die gesamte Primärsequenz codierend für die *Haematococcus pluvialis* Ketolase mit C-terminalem myc- 35 Tag, Fragment *term* (761 bp) das Polyadenylierungssignal von CaMV.

Beispiel 5: Herstellung von Expressionsvektoren zur Expression der *Haematococcus pluvialis* Ketolase in *Lycopersicon esculentum*

40

Die Expression der Ketolase aus *Haematococcus pluvialis* in *L. esculentum* erfolgte mit dem Transitpeptid *rbcS* aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle einer modifizierten Version AP3P des 45 Promoters AP3 aus *Arabidopsis thaliana* (AL132971: Nukleotidregion 9298-10200; Hill et al. (1998) Development 125: 1711-1721).

29

Das DNA Fragment, das die AP3 Promoterregion -902 bis +15 aus *Arabidopsis thaliana* beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA (nach Standardmethoden aus *Arabidopsis thaliana* isoliert) sowie der Primer PR7 (SEQ ID No. 33) und PR10 5 (SEQ ID No. 36) hergestellt.

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die das AP3-Promoterfragment 10 (-902 bis +15) beinhaltet, erfolgte in einem 50 µl Reaktionsansatz, in dem enthalten war:

- 100 ng genomischer DNA aus *A.thaliana*
- 0,25 mM dNTPs
- 15 - 0,2 mM PR7 (SEQ ID No. 33)
- 0,2 mM PR10 (SEQ ID No. 36)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0,25 µl Pfu Polymerase (Stratagene)
- 28,8 µl Aq. Dest.

20

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

	1X	94°C	2 Minuten
	35X	94°C	1 Minute
25		50°C	1 Minute
		72°C	1 Minute
	1X	72°C	10 Minuten

Das 922 Bp Amplifikat wurde unter Verwendung von Standardmethoden 30 in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pTAP3 erhalten.

Sequenzierung des Klons pTAP3 bestätigte eine Sequenz, die sich lediglich in durch eine Insertion (ein G in Position 9765 der 35 Sequenz AL132971) und einen Basenaustausch (ein G statt ein A in Position 9726 der Sequenz AL132971) von der publizierte AP3 Sequenz (AL132971, Nukleotidregion 9298-10200) unterscheidet. Diese Nukleotidunterschiede wurden in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentieren somit die 40 tatsächliche Nukleotidsequenz in den verwendeten *Arabidopsis thaliana* Pflanzen.

Die modifizierte Version AP3P wurde mittels rekombinanter PCR unter Verwendung des Plasmids pTAP3 hergestellt. Die 45 Region 10200-9771 wurde mit den Primern PR7 (SEQ ID No. 33) und Primern PR9 (SEQ ID No. 35) amplifiziert (Amplifikat A7/9),

30

die Region 9526-9285 wurde mit den PR8 (SEQ ID No. 34) und PR10 (SEQ ID No. 36) amplifiziert (Amplifikat A8/10).

Die PCR-Bedingungen waren die folgenden:

5

Die PCR-Reaktionen zur Amplifikation der DNA-Fragmente, die die Regionen Region 10200-9771 und Region 9526-9285 des AP3 Promoters beinhalten, erfolgte in 50 µl Reaktionsansätzen, in denen enthalten war:

10

- 100 ng AP3 Amplifikat (oben beschrieben)
- 0,25 mM dNTPs
- 0,2 mM sense Primer (PR7 SEQ ID No. 33 bzw. PR8 SEQ ID No. 35)
- 15 - 0,2 mM antisense Primer (PR9 SEQ ID No. 35 bzw. PR10 SEQ ID No. 36)
- 5 µl 10X PCR-Puffer (Stratagene)
- 0,25 µl Pfu Taq Polymerase (Stratagene)
- 28,8 µl Aq. Dest.

20

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

	1X	94°C	2 Minuten
	35X	94°C	1 Minute
25		50°C	1 Minute
		72°C	1 Minute
	1X	72°C	10 Minuten

Die rekombinante PCR beinhaltet Annealing der sich über eine Sequenz von 25 Nukleotiden überlappenden Amplifikate A7/9 und A8/10, Vervollständigung zu einem Doppelstrang und anschließende Amplifizierung. Dadurch entsteht eine modifizierte Version des AP3 Promoters, AP3P, in dem die Positionen 9670-9526 deletiert sind. Die Denaturierung (5 min bei 95°C) und Annealing (langsame Abkühlung bei Raumtemperatur auf 40°C) beider Amplifikate A7/9 und A8/10 erfolgte in einem 17.6 µl Reaktionsansatz, in dem enthalten war:

- 0,5 µg A7/9 Amplifikat
- 40 - 0,25 µg A8/10 Amplifikat

Das Auffüllen der 3'Enden (30 min bei 30°C) erfolgte in einem 20 µl Reaktionsansatz, in dem enthalten war:

- 45 - 17,6 µl A7/9 und A8/10-Annealingsreaktion (hergestellt wie oben beschrieben)
- 50 µM dNTPs

31

- 2 µl 1X Klenow Puffer
- 2U Klenow Enzym

Die Nukleinsäure, kodierend für die modifizierte Promoterversion
5 AP3P, wurde mittels PCR unter Verwendung eines sense spezifischen
Primers (PR7 SEQ ID No. 28) und eines antisense spezifischen Pri-
mers (PR10 SEQ ID No. 36) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

10

Die PCR zur Amplifikation des AP3P Fragmentes erfolgte in einem
50 µl Reaktionsansatz, in dem enthalten war:

- 1 µl Annealingsreaktion (hergestellt wie oben
15 beschrieben)
- 0,25 mM dNTPs
- 0,2 mM PR7 (SEQ ID No. 33)
- 0,2 mM PR10 (SEQ ID No. 36)
- 5 µl 10X PCR-Puffer (Stratagene)
- 20 - 0,25 µl Pfu Taq Polymerase (Stratagene)
- 28,8 µl Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

- | | | | |
|----|-----|------|------------|
| 25 | 1X | 94°C | 2 Minuten |
| | 35X | 94°C | 1 Minute |
| | | 50°C | 1 Minute |
| | | 72°C | 1 Minute |
| | -1X | 72°C | 10 Minuten |

30

Die PCR-Amplifikation mit SEQ ID No. 33 und SEQ ID No. 36 resul-
tierte in einem 778 Bp Fragment, das für die modifizierte Promo-
terversion AP3P kodiert. Das Amplifikat wurde in den Klonierungs-
vektor pCR2.1 (Invitrogen) kloniert und der Klon pTAP3P erhalten.

35 Sequenzierungen mit den Primern T7 und M13 bestätigten eine zur
Sequenz AL132971, Region 10200-9298 identische Sequenz, wobei die
interne Region 9285-9526 deletiert wurde. Diese Klon wurde daher
für die Klonierung in den Expressionsvektor pJIT117 (Guerineau et
al. 1988, Nucl. Acids Res. 16: 11380) verwendet.

40

Die Klonierung erfolgte durch Isolierung des 771 Bp SacI-
HindIII Fragmentes aus pTAP3P und Ligierung in den SacI-HindIII
geschnittenen Vektor pJIT117. Der Klon, der den Promoter AP3P an-
stelle des ursprünglichen Promoters d35S enthält, heißt pJAP3P.

45

32

Zur Herstellung einer Expressionskassette pJAP3PKETO2 wurde das 1027 Bp SpHI-Fragment KETO2 (in Beispiel 1 beschrieben) in den SpHI geschnittenen Vektor pJAP3P kloniert. Der Klon, der das Fragment KETO2 in der korrekten Orientierung als N-terminale Fusion mit dem *rbcS* Transitpeptid enthält, heißt pJAP3PKETO2.

Zur Herstellung einer Expressionskassette pJAP3PKETO4 wurde das 1032 Bp SpHI-EcoRI Fragment KETO4 (in Beispiel 3 beschrieben) in den SpHI-EcoRI geschnittenen Vektor pJAP3P kloniert. Der Klon, der das Fragment KETO4 in der korrekten Orientierung als N-terminale Fusion mit dem *rbcS* Transitpeptid enthält, heißt pJAP3PKETO4.

Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der AP3P-kontrollierten Ketolase aus *Haematococcus pluvialis* in *L. esculentum* erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

- Zur Herstellung des Expressionsvektors pS3AP3PKETO2 wurde das 2.8 KB bp SacI-XhoI Fragment aus pJAP3KETO2 mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert (Abbildung 8, Konstruktkarte). In der Abbildung 8 beinhaltet Fragment AP3P den modifizierten AP3P Promoter (771 bp), Fragment *rbcS* das *rbcS* Transitpeptid aus Erbse (204 bp), Fragment KETO2 (1027 bp) die gesamte Primärsequenz codierend für die *Haematococcus pluvialis* Ketolase, Fragment *term* (761 Bp) das Polyadenylierungssignal von CaMV.

- Zur Herstellung des Expressionsvektors pS3AP3PKETO4 wurde das 2.8 KB SacI-XhoI Fragment aus pJAP3PKETO4 mit dem SacI-XhoI geschnittenen Vektor pSUN3 ligiert. (Abbildung 9, Konstruktkarte). In der Abbildung 9 beinhaltet Fragment AP3P den modifizierten AP3P Promoter (771 bp), Fragment *rbcS* das *rbcS* Transitpeptid aus Erbse (204 bp), Fragment KETO4 (1038 bp) die gesamte Primärsequenz codierend für die *Haematococcus pluvialis* Ketolase mit C-terminalem myc-Tag, Fragment *term* (761 Bp) das Polyadenylierungssignal von CaMV.

Beispiel 6: Herstellung transgener *Lycopersicon esculentum* Pflanzen

Transformation und Regeneration von Tomatenpflanzen erfolgte nach der publizierten Methode von Ling und Mitarbeitern (Plant Cell Reports (1998), 17:843-847). Für die Varietät Microtom wurde mit höherer Kanamycin-Konzentration (100mg/L) selektioniert.

33

- Als Ausgangsexplantat für die Transformation dienten Kotyledonen und Hypokotyle sieben bis zehn Tage alter Keimlinge der Linie Microtom. Für die Keimung wurde das Kulturmedium nach Murashige und Skoog (1962: Murashige and Skoog, 1962, *Physiol. Plant* 15, 473-) mit 2 % Saccharose, pH 6,1 verwendet. Die Keimung fand bei 21°C bei wenig Licht (20 bis 100 µE) statt. Nach sieben bis zehn Tagen wurden die Kotyledonen quer geteilt und die Hypokotyle in ca. 5 bis 10 mm lange Abschnitte geschnitten und auf das Medium MSBN (MS, pH 6,1, 3 % Saccharose + 1 mg/l BAP, 0,1 mg/l NAA) gelegt, das am Vortag mit suspensionskultivierten Tabakzellen beschickt wurde. Die Tabakzellen wurden luftblasenfrei mit sterilem Filterpapier abgedeckt. Die Vorkultur der Explantate auf dem beschriebenen Medium erfolgte für drei bis fünf Tage. Zellen des Stammes *Agrobacterium tumefaciens* LBA4404 wurden einzeln mit den Plasmiden pS3KETO2, pS3KETO3 bzw. pS3AP3KETO2 transformiert. Von den einzelnen mit den Binaervektoren pS3KETO2, pS3KETO3 bzw. pS3KETO2 transformierten *Agrobacterium*-Stämmen wurde jeweils eine Übernachtskultur in YEB Medium mit Kanamycin (20 mg/l) bei 28 Grad Celsius kultiviert und die Zellen zentrifugiert. Das Bakterienpellet wurde mit flüssigem MS Medium (3 % Saccharose, pH 6,1) resuspendiert und auf eine optische Dichte von 0,3 (bei 600 nm) eingestellt. Die vorkultivierten Explantate wurden in die Suspension überführt und für 30 Minuten bei Zimmertemperatur unter leichtem Schütteln inkubiert. Anschließend wurden die Explantate mit sterilem Filterpapier getrocknet und für die dreitägige Co-Kultur (21°C) auf ihr Vorkulturmedium zurück gelegt.

- Nach der Co-kultur wurden die Explantate auf MSZ2 Medium (MS pH 6,1 + 3 % Saccharose, 2 mg/l Zeatin, 100 mg/l Kanamycin, 160 mg/l Timentin) transferiert und für die selektive Regeneration bei 21°C unter Schwachlicht Bedingungen (20 bis 100 µE, Lichtrhythmus 16 h / 8 h) aufbewahrt. Alle zwei bis drei Wochen erfolgte der Transfer der Explantate bis sich Sprosse bildeten. Kleine Sprosse konnten vom Explantat abgetrennt werden und auf MS (pH 6,1 + 3 % Saccharose) 160 mg/l Timentin, 30 mg/l Kanamycin, 0,1 mg/l IAA bewurzelt werden. Bewurzelte Pflanzen wurden ins Gewächshaus überführt.

- Gemäß der oben beschriebenen Transformationsmethode wurden mit folgenden Expressionskonstrukten folgende Linien erhalten:

Mit pS3KETO2 wurde erhalten: cs13-24, cs13-30, cs13-40.

Mit pS3KETO3 wurde erhalten: cs14-2, cs14-3, cs14-9, cs14-19.

- Mit pS3AP3PKETO2 wurde erhalten: cs16-15, cs16-34, cs16-35, cs16-40.

34

Beispiel 8: Charakterisierung der transgenen Früchte

Das Fruchtmaterial der transgenen Pflanzen wurde in flüssigem Stickstoff gemörst und das Pulver (etwa 250 bis 500 mg) mit 100 % Aceton extrahiert (dreimal je 500 µl). Das Lösungsmittel wurde evaporiert und die Carotinoide in 100 µl Aceton resuspendiert.

Mittels einer C30-reverse phase-Säule konnte zwischen Mono- und Diestern der Carotinoide unterschieden werden. HPLC-Laufbedingungen waren nahezu identisch mit einer publizierten Methode (Frazer et al. (2000), Plant Journal 24(4): 551-558). Eine Identifizierung der Carotinoide war aufgrund der UV-VIS-Spektren möglich.

Tabelle 1 zeigt das Carotinoidprofil in Tomatenfrüchten der gemäß der vorstehend beschriebenen Beispiele hergestellten transgenen Tomaten und Kontrolltomatenpflanzen. Im Vergleich zur genetisch nicht veränderten Kontrollpflanze weisen die genetisch veränderten Pflanzen einen Gehalt an Ketocarotinoiden und insbesondere einen Gehalt an Astaxanthin auf.

Tabelle 1

Pflanze	Lutein	Lycopin	beta-Carotin	Cryptoxanthin	Canthaxanthin	Adonirubin	Astaxanthin
Kontrolle	+	+	+	(+)	-	-	-
Kontrolle	+	+	+	(+)	-	-	-
CS13-24	-	+	+	(+)	+	+	+
CS13-30	-	+	+	(+)	+	+	+
CS13-40	-	+	+	(+)	+	+	+
CS14-2	-	+	+	(+)	+	+	+
CS14-3	-	+	+	-	+	+	+
CS14-9	-	+	+	(+)	+	+	+
CS14-19	-	+	+	-	+	+	+
CS16-15	-	+	+	(+)	+	+	+
CS16-34	-	+	+	(+)	+	+	+
CS16-35	-	+	+	-	+	+	+
CS16-40	-	+	(+)	(+)	+	+	+

+ bedeutet Carotinoid nachweisbar

- bedeutet Carotinoid nicht detektiert

(+) bedeutet Carotinoidkonzentration an der Nachweisgrenze

Abbildung 1: Biosyntheschema von Carotinoiden in Tomatenfrüchten

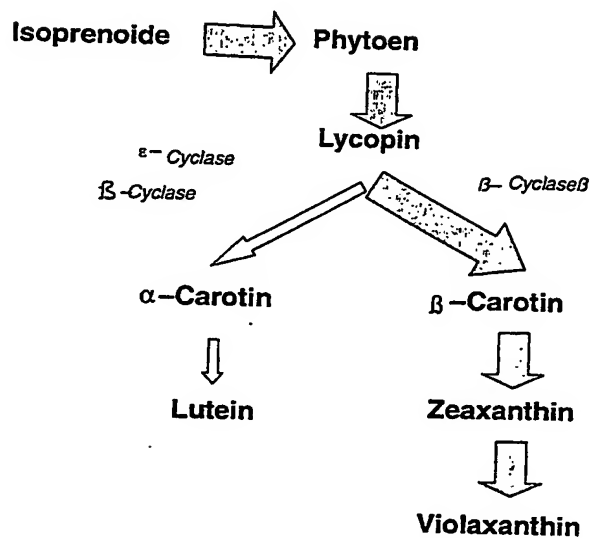


Abbildung 2: Biosyntheschema von Astaxanthin in genetisch veränderten Tomatenfrüchten

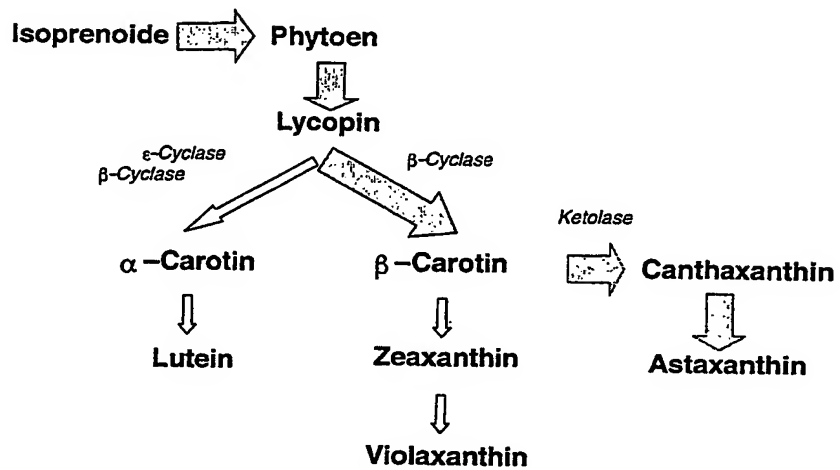


Abbildung 3: Nukleotidsequenzvergleich

KETO2.seq X86782.seq	ATTCAGCTAGCAGCGGACAGTAATGTTGGAGCAGCTTAACGGGAAGGCGCTGAGCCACTCAAGGAGAAGGAGAAGGAGGTTCAGCCAGCTCTGAGGTGTTCC ATTCAGCTAGCAGCGGACAGTAATGTTGGAGCAGCTTAACGGGAAGGCGCTGAGCCACTCAAGGAGAAGGAGAAGGAGGTTCAGCCAGCTCTGAGGTGTTCC	100 100
KETO2.seq X86782.seq	GTACATGGGCGACCCAGTACTCGCTTCGGTCAGAGGAGTCAGACCGCGGGGGGGGGGGGACTGAAGAATGCCACAAGCCACCACCTTCGGACACAAAGGG GTACATGGGCGACCCAGTACTCGCTTCGGTCAGAGGAGTCAGACCGCGGGGGGGGGGGGACTGAAGAATGCCACAAGCCACCACCTTCGGACACAAAGGG	200 200
KETO2.seq X86782.seq	CATCACAAATGGGCGTACCTGTGCATCGGCTCTGGCGGCGAGTGTTCTCCACCGCATTTTTCAAATCAAGCTTCGGAGCTCTCTGGACAGCTGCAGCTGG CATCACAAATGGGCGTACCTGTGCATCGGCTCTGGCGGCGAGTGTTCTCCACCGCATTTTTCAAATCAAGCTTCGGAGCTCTCTGGACAGCTGCAGCTGG	300 300
KETO2.seq X86782.seq	CTGCGGCGTGTGCAGATGCCACAGCTCAGCTGGTTAGCGGCAAGGAGCGAGCTGCTGCACATGTGTGTAGTAGTATCTTTGTCTGTGCTGTGACAGCGGC CTGCGGCGTGTGCAGATGCCACAGCTCAGCTGGTTAGCGGCAAGGAGCGAGCTGCTGCACATGTGTGTAGTAGTATCTTTGTCTGTGCTGTGACAGCGGC	400 400
KETO2.seq X86782.seq	TTTTTATCACACCGCATGATGCTATGCATGGCACCATGCCATGAGAAACAGGCAGCTTAATGACTTCTTGGGCAGAGTATGCATCTCTTTGTAGCGCTG TTTTTATCACACCGCATGATGCTATGCATGGCACCATGCCATGAGAAACAGGCAGCTTAATGACTTCTTGGGCAGAGTATGCATCTCTTTGTAGCGCTG	500 500
KETO2.seq X86782.seq	GTTTGATTACAACATGCTGCACCGCAAGCAATTGGAGCCACACAACACACTGGCGAGGTGGGCAAGGAACCTGACTTCACAGGGGAAACCTGGCAATT GTTTGATTACAACATGCTGCACCGCAAGCAATTGGAGCCACACAACACACTGGCGAGGTGGGCAAGGAACCTGACTTCACAGGGGAAACCTGGCAATT	600 600
KETO2.seq X86782.seq	GTGCGCTGGTTTGGACGCTTCATGTGCAGCTACATGTGCATGTGGCAGTTTGGCGGCTGGCATGGTGGAGGGGTGTCATGCAGCTGCTGGGTGGGGCAA GTGCGCTGGTTTGGACGCTTCATGTGCAGCTACATGTGCATGTGGCAGTTTGGCGGCTGGCATGGTGGAGGGGTGTCATGCAGCTGCTGGGTGGGGCAA	700 700
KETO2.seq X86782.seq	TGGCGAACCTGCTGGTGTTTCATGGGGGGGGGGGCATCTGTGGGCTTGGCTTGTCTACTTTGGCAGTACATGGGCAACAAGCTGAGCGCTGGGGC TGGCGAACCTGCTGGTGTTTCATGGGGGGGGGGGCATCTGTGGGCTTGGCTTGTCTACTTTGGCAGTACATGGGCAACAAGCTGAGCGCTGGGGC	800 800
KETO2.seq X86782.seq	CGGTCAGGCTCTTTCACCAAGGGTCATGAATCGTGGGAAGTGGGCACTAGCCAGGGTGGAGCTCGTCAGCTTTCTGAGCTGCTACCACTTGGAGCTG CGGTCAGGCTCTTTCACCAAGGGTCATGAATCGTGGGAAGTGGGCACTAGCCAGGGTGGAGCTCGTCAGCTTTCTGAGCTGCTACCACTTGGAGCTG	900 900
KETO2.seq X86782.seq	CACCTGGAGGACCAACCGCTGGGCTTTGGCGGCTGTGGGAGCTGGGCACTGGGCGGCGCTGTCTGGGCGAGTCTGGTTCTGCTGCTAG CACCTGGAGGACCAACCGCTGGGCTTTGGCGGCTGTGGGAGCTGGGCACTGGGCGGCGCTGTCTGGGCGAGTCTGGTTCTGCTGCTAG	990 990

Abbildung 4: Proteinsequenzvergleich

KETO2.pro	MQLAATVMLEQLTGSAEALKEKEKEVAGSSDVLRTWATQYSLPSEESDAA	50
X86782.pro	MQLAATVMLEQLTGSAEALKEKEKEVAGSSDVLRTWATQYSLPSEESDAA	50
KETO2.pro	RPGLKNAYKPPPSDTKGITMALAVIGSWAAVFLHAIFQIKLPTSLDQLHW	100
X86782.pro	RPGLKNAYKPPPSDTKGITMALRVIGSWAAVFLHAIFQIKLPTSLDQLHW	100
KETO2.pro	LPVSDATAQLVSGSSSLLHIVVVFFVLEFLYTGFLFITTHDAMHGHTIAMRN	150
X86782.pro	LPVSDATAQLVSGTSSLLDIVVVFFVLEFLYTGFLFITTHDAMHGHTIAMRN	150
KETO2.pro	RQLNDFLGRVCI SLYAWFDYNMLHRKHWEHHNHTGEVGKDPDFHRGNPGI	200
X86782.pro	RQLNDFLGRVCI SLYAWFDYNMLHRKHWEHHNHTGEVGKDPDFHRGNPGI	200
KETO2.pro	VPWFASFMS SYMS MWQFARLAWWT VVMQLLGAPMANLLVFMAA API LSAF	250
X86782.pro	VPWFASFMS SYMS MWQFARLAWWT VVMQLLGAPMANLLVFMAA API LSAF	250
KETO2.pro	RLFYFGTYMPHKPEPGAASGSSPAVMNWWKSRTSQASDLVSFLT CYHF DL	300
X86782.pro	RLFYFGTYMPHKPEPGAASGSSPAVMNWWKSRTSQASDLVSFLT CYHF DL	300
KETO2.pro	HWEHHRWPFAPWWELPNCRRLSGRGLVPA	329
X86782.pro	HWEHHRWPFAPWWELPNCRRLSGRGLVPA	329

Abbildung 5: Konstrukt zur Überexpression des β -C-4-Oxygenase Proteins aus *H. pluvialis* mit rbcS Transitpeptid aus Erbse unter Kontrolle des d35S-Promoters (Tomatentransformationskonstrukt)

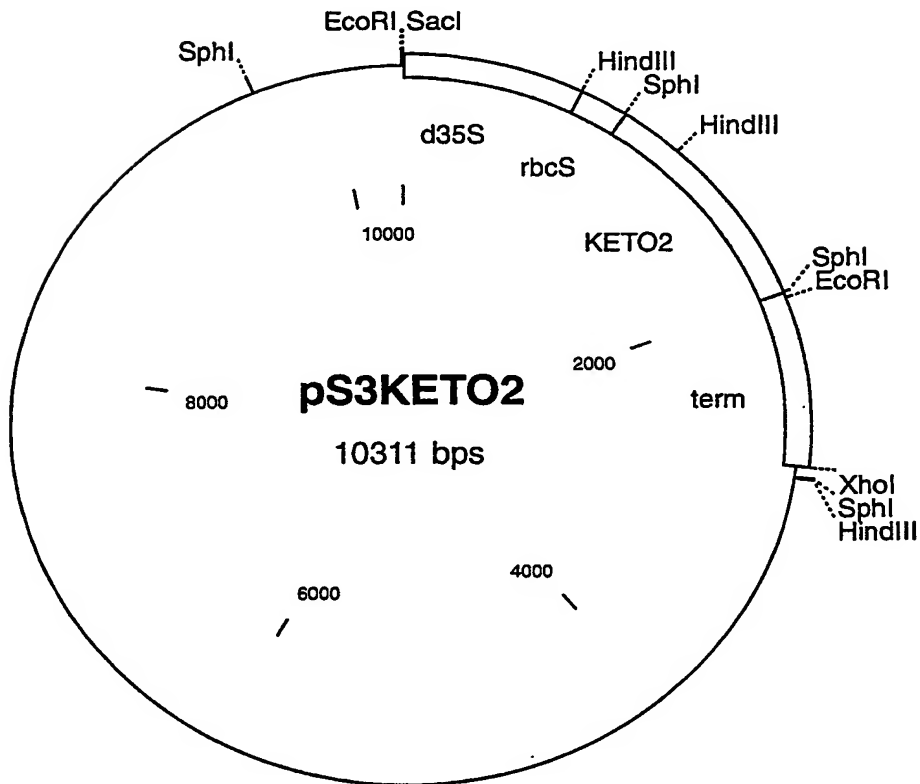


Abbildung 6: Konstrukt zur Überexpression des N-terminal verkürzten Ketolase (β -C-4-Oxygenase) Proteins aus *H. pluvialis* mit rbcS Transitpeptid aus Erbse unter Kontrolle des d35S-Promoters.

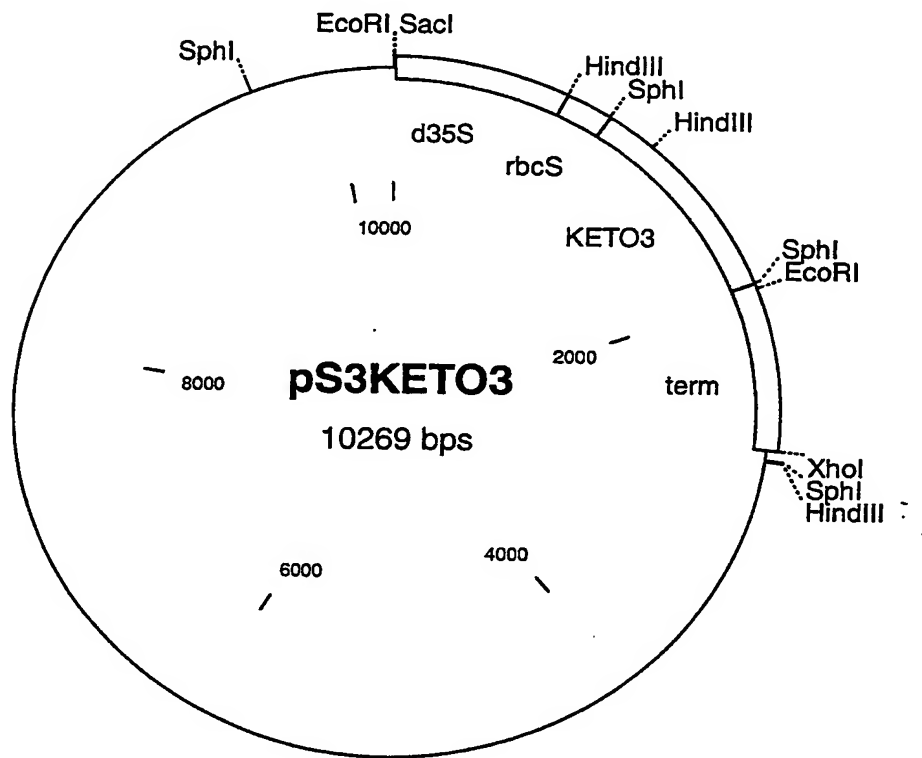


Abbildung 7: Konstrukt zur Überexpression des Ketolase (β -C-4-Oxy-
genase) Proteins aus *H. pluvialis* mit rbcS Transitpeptid aus
Erbsen und C-terminalem myc-Tag unter Kontrolle des d35S-Promo-
ters.

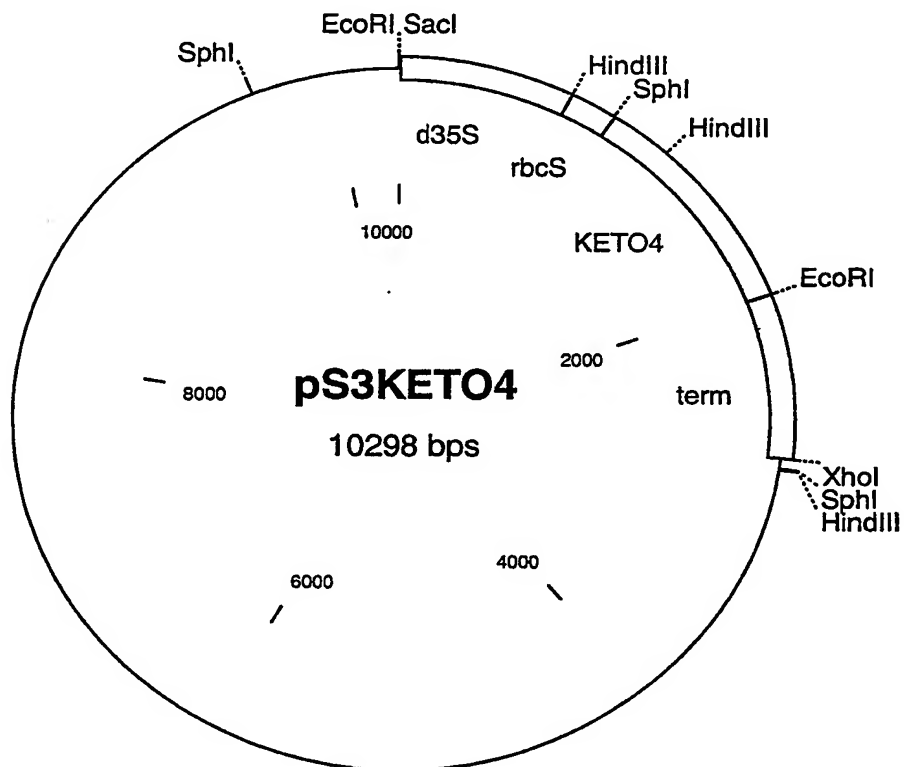


Abbildung 8: Konstrukt zur Überexpression der β -C-4-Oxygenase Protein aus *H. pluvialis* mit rbcS Transitpeptide aus Erbse unter Kontrolle des AP3P-Promoters (Tomatentransformationskonstrukt).

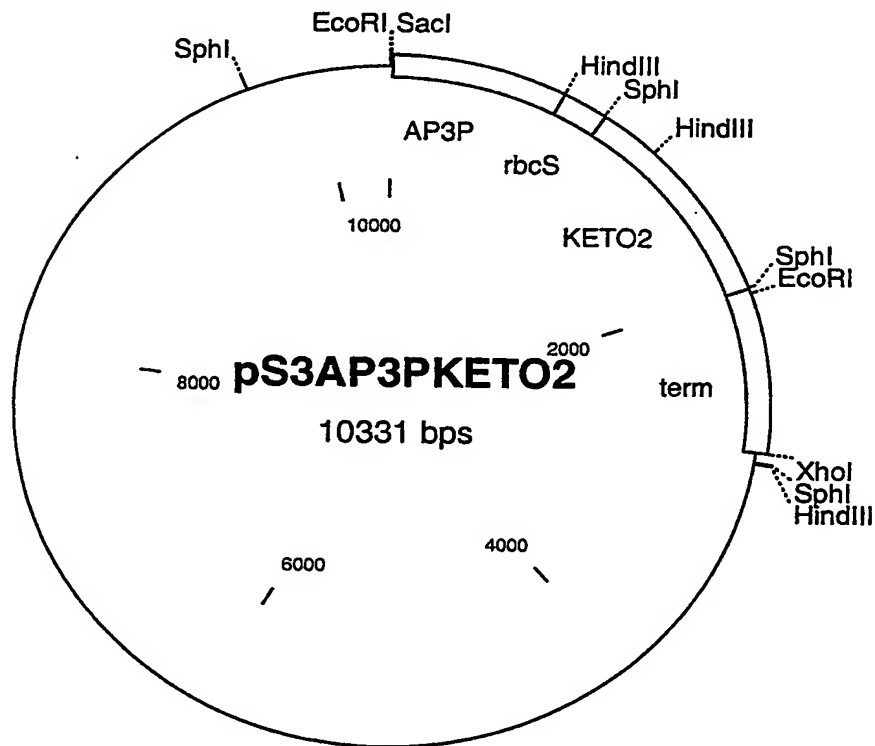
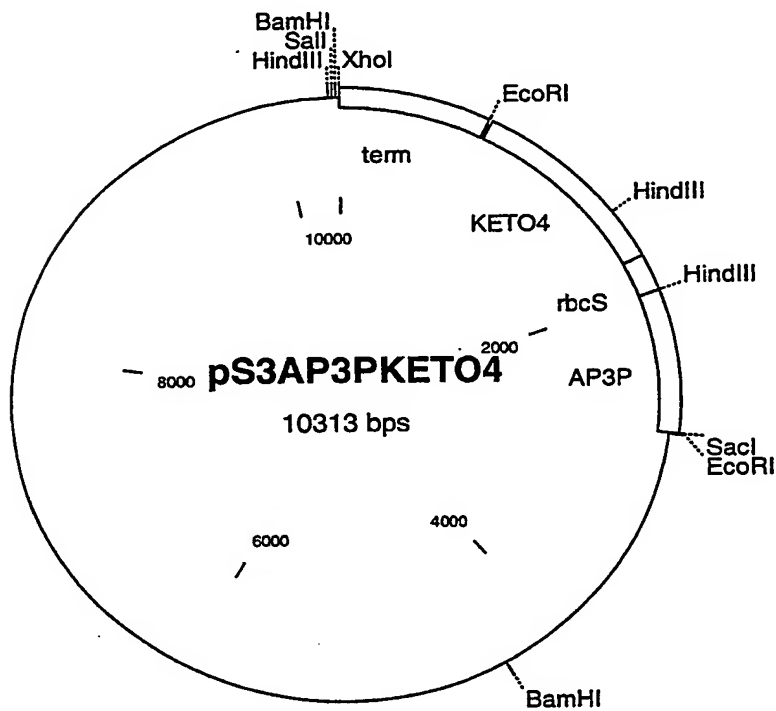


Abbildung 9: Konstrukt zur Überexpression des Ketolase (β -C-4-Oxygenase) Proteins aus *H. pluvialis* mit rbcS Transitpeptid aus Erbse und C-terminalem myc-Tag unter Kontrolle des AP3P-Promoters.



1

SEQUENCE LISTING

<110> SunGene GmbH Co. KGaA

<120> Verfahren zur herstellung von Astaxanthin in Fruechten von Pflanzen

<130> NAE 365/02

<160> 36

<170> PatentIn version 3.1

<210> 1

<211> 1771

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (166)..(1155)

<223>

<400> 1

ggcagcagct tgcacgcaag tcagcgcgcg caagtcaaca cctgccgggtc cacagcctca 60

aataataaag agctcaagcg tttgtgcgcc tcgacgtggc cagtctgcac tgccttgaac 120

ccgcgagctc cccgccgcac tgactgccat agcacagcta gacga atg cag cta gca 177
Met Gln Leu Ala
1gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag 225
Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys
5 10 15 20gag aag gag aag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg 273
Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp
25 30 35gcg acc cag tac tcg ctt ccg tca gaa gag tca gac gcg gcc cgc ccg 321
Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp Ala Ala Arg Pro
40 45 50gga ctg aag aat gcc tac aag cca cca cct tcc gac aca aag ggc atc 369
Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp Thr Lys Gly Ile
55 60 65aca atg gcg cta cgt gtc atc ggc tcc tgg gcc gca gtg ttc ctc cac 417
Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala Val Phe Leu His
70 75 80gcc att ttt caa atc aag ctt ccg acc tcc ttg gac cag ctg cac tgg 465
Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp Gln Leu His Trp
85 90 95 100

ctg ccc gtg tca gat gcc aca gct cag ctg gtt agc ggc acg agc agc 513

Leu	Pro	Val	Ser	Asp	Ala	Thr	Ala	Gln	Leu	Val	Ser	Gly	Thr	Ser	Ser		
				105					110					115			
ctg	ctc	gac	atc	gtc	gta	gta	ttc	ttt	gtc	ctg	gag	ttc	ctg	tac	aca	561	
Leu	Leu	Asp	Ile	Val	Val	Val	Phe	Phe	Val	Leu	Glu	Phe	Leu	Tyr	Thr		
				120					125					130			
ggc	ctt	ttt	atc	acc	acg	cat	gat	gct	atg	cat	ggc	acc	atc	gcc	atg	609	
Gly	Leu	Phe	Ile	Thr	Thr	His	Asp	Ala	Met	His	Gly	Thr	Ile	Ala	Met		
				135					140					145			
aga	aac	agg	cag	ctt	aat	gac	ttc	ttg	ggc	aga	gta	tgc	atc	tcc	ttg	657	
Arg	Asn	Arg	Gln	Leu	Asn	Asp	Phe	Leu	Gly	Arg	Val	Cys	Ile	Ser	Leu		
				150					155					160			
tac	gcc	tgg	ttt	gat	tac	aac	atg	ctg	cac	cgc	aag	cat	tgg	gag	cac	705	
Tyr	Ala	Trp	Phe	Asp	Tyr	Asn	Met	Leu	His	Arg	Lys	His	Trp	Glu	His		
				165					170					175			180
cac	aac	cac	act	ggc	gag	gtg	ggc	aag	gac	cct	gac	ttc	cac	agg	gga	753	
His	Asn	His	Thr	Gly	Glu	Val	Gly	Lys	Asp	Pro	Asp	Phe	His	Arg	Gly		
				185					190					195			
aac	cct	ggc	att	gtg	ccc	tgg	ttt	gcc	agc	ttc	atg	tcc	agc	tac	atg	801	
Asn	Pro	Gly	Ile	Val	Pro	Trp	Phe	Ala	Ser	Phe	Met	Ser	Ser	Tyr	Met		
				200					205					210			
tcg	atg	tgg	cag	ttt	gcg	cgc	ctc	gca	tgg	tgg	acg	gtg	gtc	atg	cag	849	
Ser	Met	Trp	Gln	Phe	Ala	Arg	Leu	Ala	Trp	Trp	Thr	Val	Val	Met	Gln		
				215					220					225			
ctg	ctg	ggc	gcg	cca	atg	gcg	aac	ctg	ctg	gtg	ttc	atg	gcg	gcc	gcg	897	
Leu	Leu	Gly	Ala	Pro	Met	Ala	Asn	Leu	Leu	Val	Phe	Met	Ala	Ala	Ala		
				230					235					240			
ccc	atc	ctg	tcc	gcc	ttc	cgc	ttg	ttc	tac	ttt	ggc	acg	tac	atg	ccc	945	
Pro	Ile	Leu	Ser	Ala	Phe	Arg	Leu	Phe	Tyr	Phe	Gly	Thr	Tyr	Met	Pro		
				245					250					255			260
cac	aag	cct	gag	cct	ggc	gcc	gcg	tca	ggc	tct	tca	cca	gcc	gtc	atg	993	
His	Lys	Pro	Glu	Pro	Gly	Ala	Ala	Ser	Gly	Ser	Ser	Pro	Ala	Val	Met		
				265					270					275			
aac	tgg	tgg	aag	tcg	cgc	act	agc	cag	gcg	tcc	gac	ctg	gtc	agc	ttt	1041	
Asn	Trp	Trp	Lys	Ser	Arg	Thr	Ser	Gln	Ala	Ser	Asp	Leu	Val	Ser	Phe		
				280					285					290			
ctg	acc	tgc	tac	cac	ttc	gac	ctg	cac	tgg	gag	cac	cac	cgc	tgg	ccc	1089	
Leu	Thr	Cys	Tyr	His	Phe	Asp	Leu	His	Trp	Glu	His	His	Arg	Trp	Pro		
				295					300					305			
ttc	gcc	ccc	tgg	tgg	gag	ctg	ccc	aac	tgc	cgc	cgc	ctg	tct	ggc	cga	1137	
Phe	Ala	Pro	Trp	Trp	Glu	Leu	Pro	Asn	Cys	Arg	Arg	Leu	Ser	Gly	Arg		
				310					315					320			
ggc	ctg	gtt	cct	gcc	tag	ctgg	acac	acac	tg	cagt	gggc	cctg	ctg	cca		1	

325

gctgggcatg caggttgtgg caggactggg tgaggtgaaa agctgcaggc gctgctgccg 1245
gacacgctgc atgggctacc ctgtgtagct gccgccacta ggggaggggg tttgtagctg 1305
tcgagcttgc cccatggatg aagctgtgta gtggtgcagg gagtacaccc acaggccaac 1365
acccttgcag gagatgtctt gcgtcgggag gagtggtggg cagtgtagat gctatgattg 1425
tatcttaatg ctgaagcctt taggggagcg acacttagtg ctgggcaggc aacgccctgc 1485
aaggtgcagg cacaagctag gctggacgag gactcgggtg caggcagggtg aagaggtgcg 1545
ggaggggtggt gccacaccca ctgggcaaga ccatgctgca atgctggcgg tgtggcagtg 1605
agagctgcgt gattaactgg gctatggatt gtttgagcag tctcacttat tctttgatat 1665
agatactggt caggcaggtc aggagagtga gtatgaacaa gttgagaggt ggtgcgctgc 1725
ccctgcgctt atgaagctgt aacaataaag tggttcaaaa aaaaaa 1771

<210> 2

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 2

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala
1 5 10 15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val
20 25 30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp
35 40 45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp
50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala
65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp
85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser
100 105 110

4

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr
210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser
260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp
275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His
290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg
305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala
325

5

<210> 3
 <211> 1662
 <212> DNA
 <213> Haematococcus pluvialis

<220>
 <221> CDS
 <222> (168)..(1130)
 <223>

<400> 3
 cgggggcaact caagaaattc aacagctgca agcgcgcccc agcctcacag cgccaagtga 60
 gctatcgacg tggttgtgag cgctcgacgt ggtccactga cgggcctgtg agcctctgcg 120
 ctccgtcctc tgccaaatct cgcgtcgggg cctgcctaag tcgaaga atg cac gtc 176
 Met His Val
 1
 gca tcg gca cta atg gtc gag cag aaa ggc agt gag gca gct gct tcc 224
 Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser
 5 10 15
 agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc 272
 Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser
 20 25 30 35
 gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct 320
 Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala Tyr Lys Pro
 40 45 50
 cca gca tct gac gcc aag ggc atc acg atg gcg ctg acc atc att ggc 368
 Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr Ile Ile Gly
 55 60 65
 acc tgg acc gca gtg ttt tta cac gca ata ttt caa atc agg cta ccg 416
 Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile Arg Leu Pro
 70 75 80
 aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa gcc aca gcc 464
 Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu Ala Thr Ala
 85 90 95
 cag ctt ttg ggc gga agc agc agc cta ctg cac atc gct gca gtc ttc 512
 Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala Ala Val Phe
 100 105 110 115
 att gta ctt gag ttc ctg tac act ggt cta ttc atc acc aca cat gac 560
 Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp
 120 125 130
 gca atg cat ggc acc ata gct ttg agg cac agg cag ctc aat gat ctc 608
 Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu
 135 140 145
 ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac tac agc atg 656

6

Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Ser Met	
150 155 160	
ctg cat cgc aag cac tgg gag cac cac aac cat act ggc gaa gtg ggg	704
Leu His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly	
165 170 175	
aaa gac cct gac ttc cac aag gga aat ccc ggc ctt gtc ccc tgg ttc	752
Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe	
180 185 190 195	
gcc agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg	800
Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu	
200 205 210	
gca tgg tgg gca gtg gtg atg caa atg ctg ggg gcg ccc atg gca aat	848
Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro Met Ala Asn	
215 220 225	
ctc cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc	896
Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu	
230 235 240	
ttc tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca	944
Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala	
245 250 255	
gca ggc tct cag gtg atg gcc tgg ttc agg gcc aag aca agt gag gca	992
Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr Ser Glu Ala	
260 265 270 275	
tct gat gtg atg agt ttc ctg aca tgc tac cac ttt gac ctg cac tgg	1040
Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp	
280 285 290	
gag cac cac agg tgg ccc ttt gcc ccc tgg tgg cag ctg ccc cac tgc	1088
Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu Pro His Cys	
295 300 305	
cgc cgc ctg tcc ggg cgt ggc ctg gtg cct gcc ttg gca tga	1130
Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala	
310 315 320	
cctgggtccct ccgctggtga cccagcgtct gcacaagagt gtcattgctac aggggtgctgc	1190
ggccagtggc agcgcagtgc actctcagcc tgtatggggc taccgctgtg ccaactgagca	1250
ctgggcatgc cactgagcac tgggcgtgct actgagcaat gggcgtgcta ctgagcaatg	1310
ggcgtgctac tgacaatggg cgtgctactg gggctctggca gtggctagga tggagtttga	1370
tgcattcagt agcgggtggc aacgtcatgt ggatgggtgga agtgctgagg ggtttaggca	1430
gccggcattt gagaggggcta agttataaat cgcattgctgc tcatgcgcac atatctgcac	1490
acagccaggg aaatcccttc gagagtgatt atgggacact tgtattgggtt tcgtgctatt	1550

7

gttttattca gcagcagtagc ttagtgaggg tgagagcagg gtggtgagag tggagtgagt 1610

gagtatgaac ctggtcagcg aggtgaacag cctgtaatga atgactctgt ct 1662

<210> 4

<211> 320

<212> PRT

<213> Haematococcus pluvialis

<400> 4

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala
1 5 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His
20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala
35 40 45

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr
50 55 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile
65 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu
85 90 95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala
100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr
115 120 125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu
130 135 140

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp
145 150 155 160

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly
165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val

8

180

185

190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe
 195 200 205

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro
 210 215 220

Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala
 225 230 235 240

Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro
 245 250 255

Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr
 260 265 270

Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp
 275 280 285

Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu
 290 295 300

Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala
 305 310 315 320

<210> 5

<211> 729

<212> DNA

<213> Agrobacterium aurantiacum

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 5

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg 48
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat 96
 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
 20 25 30

gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca 144
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala

35	40	45	
aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg			192
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala			
50	55	60	
cat gac gcg atg cac ggg tcg gtg gtg ccg ggg cgt ccg cgc gcc aat			240
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn			
65	70	75	80
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg			288
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp			
	85	90	95
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc			336
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr			
	100	105	110
gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc			384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala			
	115	120	125
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc			432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro			
	130	135	140
gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac			480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr			
	145	150	155
gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gcg tcg atc cag ctg ttc			528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe			
	165	170	175
gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg			576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro			
	180	185	190
gac cgc cac aat gcg cgg tcg tcg cgg atc agc gac ccc gtg tcg ctg			624
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu			
	195	200	205
ctg acc tgc ttt cac ttt ggc ggt tat cat cac gaa cac cac ctg cac			672
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His			
	210	215	220
ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac			720
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp			
	225	230	235
acc gca tga			729
Thr Ala			

<210> 6

<211> 242

10

<212> PRT

<213> Agrobacterium aurantiacum

<400> 6

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

11

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
 225 230 235 240

Thr Ala

<210> 7
 <211> 1631
 <212> DNA
 <213> Alcaligenes sp.

<220>
 <221> CDS
 <222> (99)..(827)
 <223>

<400> 7
 ctgcaggccg ggcccgggtgg ccaatgggtcg caaccggcag gactggaaca ggacggcggg 60
 ccggtctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct 116
 Met Ser Gly Arg Lys Pro
 1 5
 ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc 164
 Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile
 10 15 20
 ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat 212
 Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp
 25 30 35
 gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc 260
 Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr
 40 45 50
 tgg ctg tgc gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg 308
 Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly
 55 60 65 70
 tcc gtg gtg ccg ggg cgg ccg cgc gcc aat gcg gcg atc ggg caa ctg 356
 Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu
 75 80 85
 gcg ctg tgg ctc tat gcg ggg ttc tgc tgg ccc aag ctg atc gcc aag 404
 Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys
 90 95 100
 cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc 452
 His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe
 105 110 115

12

ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat	500
Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr	
120 125 130	
ttc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat	548
Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro Val Ile Val Thr Thr Tyr	
135 140 145 150	
gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc	596
Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val	
155 160 165	
ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg	644
Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu	
170 175 180	
ccc cac cgc ccg gga cat gac gat ttt ccc gac cgg cac aac gcg agg	692
Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg	
185 190 195	
tcg acc ggc atc ggc gac ccg ttg tca cta ctg acc tgc ttc cat ttc	740
Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe	
200 205 210	
ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg	788
Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp	
215 220 225 230	
cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct	837
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala	
235 240	
cattgtcgtg ggcacagtc tctgtgatgga gctgaccgcc tattccgtcc accgctggat	897
tatgcacggc cccctaggct ggggctggca caagtcccat cacgaagagc acgaccacgc	957
gttggagaag aacgacctct acggcgtcgt cttcgcggtg ctggcgacga tcctcttcac	1017
cgtgggcgcc tattggtggc cgggtgctgtg gtggatcgcc ctgggcatga cggctctatgg	1077
gttgatctat ttcatacctgc acgacgggct tgtgcatcaa cgctggccgt ttcggtatat	1137
tccgcggcgg ggctatttcc gcaggctcta ccaagctcat cgctgcacc acgcggtcga	1197
ggggcgggac cactgcgtca gcttcggcct catctatgcc ccaccgtgg acaagctgaa	1257
gcaggatctg aagcggtcgg gtgtcctgcg ccccgaggac gagcgtccgt cgtgatctct	1317
gatcccgggc tggccgcatg aaatccgacg tgctgctggc aggggcgggc cttgccaaag	1377
gactgatcgc gctggcgatc cgcaaggcgc ggcccgacct tcgctgctg ctgctggacc	1437
gtgcggcggg cgctcggac gggcatactt ggtcctgcca cgacaccgat ttggcgccgc	1497
actggctgga ccgcctgaag ccgatcaggc gtggcgactg gcccgatcag gaggtgcggt	1557

13

tcccagacca ttcgcgaagg ctccgggccc gatatggctc gatcgacggg cgggggctga 1617
tgcgtgcggt gacc 1631

<210> 8
<211> 242
<212> PRT
<213> Alcaligenes sp.

<400> 8

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu
1 5 10 15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe
20 25 30

Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu
35 40 45

Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr
100 105 110

Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly
115 120 125

Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro

14

180

185

190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly
 225 230 235 240

Arg Ala

<210> 9
 <211> 729
 <212> DNA
 <213> Paracoccus marcusii

<220>
 <221> CDS
 <222> (1)..(729)
 <223>

<400> 9
 atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc aca agc ctg 48
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtg cat 96
 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
 20 25 30

gcg ctg tgg ttt ctg gac gcg gcg gcc cat ccc atc ctg gcg gtc gcg 144
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
 35 40 45

aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192
 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
 50 55 60

cat gac gcg atg cac ggg tcg gtc gtg ccg ggg cgt ccg cgc gcc aat 240
 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
 65 70 75 80

gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg 288
 Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
 85 90 95

cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc 336
 Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr

15

100	105	110	
gac gac gac cca gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc			384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala			
115	120	125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc			432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro			
130	135	140	
gtc atc gtg acg gtc tat gcg ctg atc ctg ggg gat cgc tgg atg tac			480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr			
145	150	155	160
gtg gtc ttc tgg ccg ttg ccg tcg atc ctg gcg tcg atc cag ctg ttc			528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe			
165	170	175	
gtg ttc ggc act tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg			576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro			
180	185	190	
gac cgc cat aat gcg ccg tcg tcg ccg atc agc gac cct gtg tcg ctg			624
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu			
195	200	205	
ctg acc tgc ttt cat ttt ggc ggt tat cat cac gaa cac cac ctg cac			672
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His			
210	215	220	
ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac			720
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp			
225	230	235	240
acc gca tga			729
Thr Ala			

<210> 10
 <211> 242
 <212> PRT
 <213> *Paracoccus marcusii*

<400> 10

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu			
1	5	10	15
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His			
20	25	30	
Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala			
35	40	45	

16

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

Thr Ala

<210> 11

<211> 1629

<212> DNA

<213> Synechococcus sp.

<220>

<221> CDS

<222> (1)..(1629)

<223>

<400> 11

atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta	48
Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu	
1 5 10 15	

gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta	96
Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu	
20 25 30	

gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg	144
Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met	
35 40 45	

ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac	192
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His	
50 55 60	

gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag	240
Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln	
65 70 75 80	

tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg	288
Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly	
85 90 95	

ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt	336
Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys	
100 105 110	

gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa	384
Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln	
115 120 125	

ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt	432
Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe	
130 135 140	

aat gct ccg ccc cag gct tta cta gat tta gcc ctg aac tat ggt tgg	480
Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp	
145 150 155 160	

gaa aac tta aaa tcc gtg ctg gcg atc gcc ggg tcg aaa acc aag gcg	528
Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala	
165 170 175	

ttg gat ttt atc cgc act atg atc ggc tcc ccg gaa gat gtg ctc aat	576
Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn	
180 185 190	

gaa tgg ttc gac agc gaa cgg gtt aaa gct cct tta gct aga cta tgt	624
---	-----

18

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys	
195	200 205
tcg gaa att ggc gct ccc cca tcc caa aag ggt agt agc tcc ggc atg	672
Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met	
210	215 220
atg atg gtg gcc atg cgg cat ttg gag gga att gcc aga cca aaa gga	720
Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly	
225	230 235 240
ggc act gga gcc ctc aca gaa gcc ttg gtg aag tta gtg caa gcc caa	768
Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln	
	245 250 255
ggg gga aaa atc ctc act gac caa acc gtc aaa cgg gta ttg gtg gaa	816
Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu	
	260 265 270
aac aac cag gcg atc ggg gtg gag gta gct aac gga gaa cag tac cgg	864
Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg	
	275 280 285
gcc aaa aaa ggc gtg att tct aac atc gat gcc cgc cgt tta ttt ttg	912
Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu	
	290 295 300
caa ttg gtg gaa ccg ggg gcc cta gcc aag gtg aat caa aac cta ggg	960
Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly	
305	310 315 320
gaa cga ctg gaa cgg cgc act gtg aac aat aac gaa gcc att tta aaa	1008
Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys	
	325 330 335
atc gat tgt gcc ctc tcc ggt tta ccc cac ttc act gcc atg gcc ggg	1056
Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly	
	340 345 350
ccg gag gat cta acg gga act att ttg att gcc gac tcg gta cgc cat	1104
Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His	
	355 360 365
gtc gag gaa gcc cac gcc ctc att gcc ttg ggg caa att ccc gat gct	1152
Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala	
	370 375 380
aat ccg tct tta tat ttg gat att ccc act gta ttg gac ccc acc atg	1200
Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met	
385	390 395 400
gcc ccc cct ggg cag cac acc ctc tgg atc gaa ttt ttt gcc ccc tac	1248
Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr	
	405 410 415
cgc atc gcc ggg ttg gaa ggg aca ggg tta atg ggc aca ggt tgg acc	1296
Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr	

19

420	425	430	
gat gag tta aag gaa aaa gtg gcg gat cgg gtg att gat aaa tta acg			1344
Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr			
435	440	445	
gac tat gcc cct aac cta aaa tct ctg atc att ggt cgc cga gtg gaa			1392
Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu			
450	455	460	
agt ccc gcc gaa ctg gcc caa cgg ctg gga agt tac aac ggc aat gtc			1440
Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val			
465	470	475	480
tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta			1488
Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu			
485	490	495	
ccg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca			1536
Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr			
500	505	510	
ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga			1584
Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg			
515	520	525	
aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa			1629
Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp			
530	535	540	

<210> 12
<211> 542
<212> PRT
<213> Synechococcus sp.

<400> 12

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu	
1 5 10 15	
Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu	
20 25 30	
Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met	
35 40 45	
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His	
50 55 60	
Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln	
65 70 75 80	

20

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly
85 90 95

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys
100 105 110

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe
130 135 140

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp
145 150 155 160

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala
165 170 175

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn
180 185 190

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys
195 200 205

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met
210 215 220

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly
225 230 235 240

Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln
245 250 255

Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu
260 265 270

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg
275 280 285

Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu
290 295 300

21

Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly
305 310 315 320

Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys
325 330 335

Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly
340 345 350

Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His
355 360 365

Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala
370 375 380

Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met
385 390 395 400

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr
405 410 415

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr
420 425 430

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr
435 440 445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu
450 455 460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
465 470 475 480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
485 490 495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
500 505 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp

22

530

535

540

<210> 13
 <211> 776
 <212> DNA
 <213> Bradyrhizobium sp.

<220>
 <221> CDS
 <222> (1)..(774)
 <223>

<400> 13
 atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cgg cgc 48
 Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg
 1 5 10 15

gac gat gcg agg cag cgc cgc gtc ggt ctc acg ctg gcc gcg gtc atc 96
 Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile
 20 25 30

atc gcc gcc tgg ctg gtg ctg cat gtc ggt ctg atg ttc ttc tgg ccg 144
 Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro
 35 40 45

ctg acc ctt cac agc ctg ctg ccg gct ttg cct ctg gtg gtg ctg cag 192
 Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln
 50 55 60

acc tgg ctc tat gta ggc ctg ttc atc atc gcg cat gac tgc atg cac 240
 Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His
 65 70 75 80

ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga cag 288
 Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln
 85 90 95

ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc 336
 Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val
 100 105 110

gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat 384
 Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp
 115 120 125

ttc gac gag gtg ccg ccg cac ggc ttc tgg cac tgg ttc gcc agc ttt 432
 Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe
 130 135 140

ttc ctg cac tat ttc ggc tgg aag cag gtc gcg atc atc gca gcc gtc 480
 Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
 145 150 155 160

tcg ctg gtt tat cag ctc gtc ttc gcc gtt ccc ttg cag aac atc ctg 528
 Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu

23

165										170					175					
ctg	ttc	tgg	gcg	ctg	ccc	ggg	ctg	ctg	tcg	gcg	ctg	cag	ctg	ttc	acc	576				
Leu	Phe	Trp	Ala	Leu	Pro	Gly	Leu	Leu	Ser	Ala	Leu	Gln	Leu	Phe	Thr					
180						185					190									
ttc	ggc	acc	tat	ctg	ccg	cac	aag	ccg	gcc	acg	cag	ccc	ttc	gcc	gat	624				
Phe	Gly	Thr	Tyr	Leu	Pro	His	Lys	Pro	Ala	Thr	Gln	Pro	Phe	Ala	Asp					
195						200					205									
cgc	cac	aac	gcg	cgg	acg	agc	gaa	ttt	ccc	gcg	tgg	ctg	tcg	ctg	ctg	672				
Arg	His	Asn	Ala	Arg	Thr	Ser	Glu	Phe	Pro	Ala	Trp	Leu	Ser	Leu	Leu					
210						215					220									
acc	tgc	ttc	cac	ttc	ggc	ttt	cat	cac	gag	cat	cat	ctg	cat	ccc	gat	720				
Thr	Cys	Phe	His	Phe	Gly	Phe	His	His	Glu	His	His	Leu	His	Pro	Asp					
225						230					235				240					
gcg	ccg	tgg	tgg	cgg	ctg	ccg	gag	atc	aag	cgg	cgg	gcc	ctg	gaa	agg	768				
Ala	Pro	Trp	Trp	Arg	Leu	Pro	Glu	Ile	Lys	Arg	Arg	Ala	Leu	Glu	Arg					
245						250					255									
cgt	gac	ta													776					
Arg	Asp																			

<210> 14
 <211> 258
 <212> PRT
 <213> Bradyrhizobium sp.

<400> 14

Met	His	Ala	Ala	Thr	Ala	Lys	Ala	Thr	Glu	Phe	Gly	Ala	Ser	Arg	Arg
1				5					10					15	

Asp	Asp	Ala	Arg	Gln	Arg	Arg	Val	Gly	Leu	Thr	Leu	Ala	Ala	Val	Ile
		20						25					30		

Ile	Ala	Ala	Trp	Leu	Val	Leu	His	Val	Gly	Leu	Met	Phe	Phe	Trp	Pro
		35					40					45			

Leu	Thr	Leu	His	Ser	Leu	Leu	Pro	Ala	Leu	Pro	Leu	Val	Val	Leu	Gln
	50					55					60				

Thr	Trp	Leu	Tyr	Val	Gly	Leu	Phe	Ile	Ile	Ala	His	Asp	Cys	Met	His
65					70					75					80

Gly	Ser	Leu	Val	Pro	Phe	Lys	Pro	Gln	Val	Asn	Arg	Arg	Ile	Gly	Gln
				85					90					95	

24

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val
100 105 110

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp
115 120 125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe
130 135 140

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
145 150 155 160

Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu
165 170 175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
245 250 255

Arg Asp

<210> 15
<211> 777
<212> DNA
<213> Nostoc sp.

<220>
<221> CDS
<222> (1)..(777)
<223>

<400> 15

25

atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta	48
Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu	
1 5 10 15	
ttg tca tcg aca atc aga gat gat aaa aat att aat aag ggt ata ttt	96
Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe	
20 25 30	
att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta	144
Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu	
35 40 45	
ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc	192
Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala	
50 55 60	
atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat	240
Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His	
65 70 75 80	
gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat	288
Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn	
85 90 95	
ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa	336
Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys	
100 105 110	
gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat	384
Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp	
115 120 125	
tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg	432
Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp	
130 135 140	
tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga	480
Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly	
145 150 155 160	
tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa	528
Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu	
165 170 175	
aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta	576
Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val	
180 185 190	
caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt	624
Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly	
195 200 205	
ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt	672
Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe	
210 215 220	
tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac	720

26

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata 768
 Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

tct tta taa 777
 Ser Leu

<210> 16
 <211> 258
 <212> PRT
 <213> Nostoc sp.

<400> 16

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
 1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe
 20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu
 35 40 45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala
 50 55 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
 65 70 75 80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn
 85 90 95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys
 100 105 110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp
 115 120 125

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp
 130 135 140

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly
 145 150 155 160

27

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu
165 170 175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
180 185 190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
195 200 205

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
245 250 255

Ser Leu

<210> 17
<211> 2093
<212> DNA
<213> Tomato

<220>
<221> promoter
<222> (1)..(2093)
<223>

<400> 17
tttgccagta ttacaacagc ttatatgttg agcaggtaaa agcttcaatg ccctattctt 60
tctacagtta tcaatgttgc tcgtctaata tctggtgttc ttctcgaaat gtcaattggc 120
ttgcagcaca ttgtcctcta atatccattc aagcttctta gatgatgaaa catttgtcaa 180
atttattaat ttcatagtgt tcagtctcaa ttcttttagct gggttcctcat agtaaagttg 240
tctaatatga aatgaaaatg ttctgtgtgt tgtactaata ccttttcatg gttgtctata 300
gaacgtcgat gaagagccaa acagaaacta ttttgggctg cgatttctga taccattgta 360
tctgaatgct ggggtgggagc tcatcagaag ctttacaatg ggtcacatat atggagccgg 420
tatgaggaat gctgggaatc agttgcgttt cgcgtgctag gacttttcct tcctgggtatt 480

28

tctgcccaca gccagttga ttacgtgaac tccgtcagac ttggaaagga gagaagtacc 540
caaatgtcgt ctttttagaa atacttttgt cacaaaatag cggggtttac agctacagaa 600
gatcatgcag aaggcgtcca gtttagtttt tgaaggttgt ttggagttta tttatctaaa 660
gtaaacttaa atcagctttt tgtttatgag ttcagtgaac tatatgttca aataagactt 720
ccctttgtag atatgtgttt tttttgttgt tgagcacttt gtgtgcattg gataaacccc 780
caacgtgtaa tagctaccat acaagagaag taactcgcac tgtccatgtc ttatgtggct 840
cgactcagaa agcattcagg gggattgata accaccctcc aaaccaactg aaccattgtg 900
aataaccacc cttcaaatca accgagtcct cgtgaaggac aaatatgtgg ttttatatac 960
attaaatttt gtttttacat gcttcctctt acttcttttag tttctttgac catatcttgc 1020
gtttttccct tctgtaattg acacttttct tcaaaccatc cagcaatgtg gaagcttgac 1080
gattttcctt cagagtagaa attgaaaaga atcaactaaa aaggatagtc cttcgatttg 1140
atctccggct taaaaataaa ctaataagaa tgagagagcg aataatagaa tattttgaaa 1200
ttttaagat attcaactat gttaaattgc gttataaatt tcttaaatta gtagcaccta 1260
atagtttagt tctcaaaagt caaaactact acataatgtg ctcatTTTTc acattaaaat 1320
gcctacatga tgtaaaagta aaactcgtag cattctacgt gttttactca actcaaacat 1380
cctgttcatt ttaataaacg tacgatgagc ttctctctcc aattttcttt tctttttttt 1440
ttttaaaaaa atattttttt ttatatcaat ccaaatgggc tccaatttat cataaattag 1500
gtagaaactt agatattaaa gaaagaaaag ggtttatctc gcaagtgtgg ctatgggtggg 1560
acgtgtcaaa ttttggattg tagccaaaca tgagatttga tttaaaggga attggccaaa 1620
tcaccgaaag caggcatctt catcataaat tagtttgttt atttatacag aattatacgc 1680
ttttactagt tatagcattc ggtatctttt tctgggtaac tgccaaacca ccacaaattt 1740
caagtttcca ttttaactct caacttcaac ccaaccaaatt tttttgctt aattgtgcag 1800
aaccactccc tatatcttct aggtgctttc attcgttccg aggtaagaaa agatttttgt 1860
ttctttgaat gctttatgcc actcgtttaa cttctgaggt ttgtggatct tttaggcgac 1920
tttttttttt tttgtatgta aaatttgttt cataaatgct tctcaacata aatcttgaca 1980
aagagaagga attttaccaa gtatttaggt tcagaaatgg ataattttct tactgtgaaa 2040
tatccttatg gcaggtttta ctgttatttt tcagtaaaat gcctcaaatt gga 2093

<211> 4760
<212> DNA
<213> Tomato

<220>
<221> promoter
<222> (1)..(4760)
<223>

<400> 18
tctagattga aataaacctt attgcattta gtatatgaga atgcatctat aaaataatgt 60
ctatTTTTTgg tggaaaatat ttgtgcgcca aagcacgggt tgtatTTTtat atTTTacaat 120
atTTTtgcac ggtaatatag ttgcaagggt ttacaaacga attatctctt gaactTTTaaa 180
ttaagttcac agTTTattcc aaaaataatg ttcaacttct aatcatatct cccctatttg 240
ctagaaaaat ataacattta cgcccaactt catttaggat ccattTTTtat gcatgggtgga 300
gcaattggat catatactac atatTTTTTTT aaaaaaata gatagaaatt atttaatctt 360
gattccgaat caattgtgat gggaaaacct tattagTTtg atgtgtacat ataatgTTTT 420
atgtcaaata aatttatttt atactaaatt ttatttgaaa gtatTTTTtct cataacaaat, 480
aatttaacta tattggagac atgaaaattc tacaaaacca acttgcatta tcaacataat 540
tttatagttt gaaattgtgc tcttaattaa acaattcaag ataacaatct ggtaaaatta 600
aaattacaag ttgataacaa acatatacat atgtacatct catagatgca ttcattaaat 660
catataatag taaatgcttc acaatagaag ggtctatatt cattTTTTTTT ttatgtgtca 720
aacaatTTtg aggaattcaa tttcatcttt aactggtaca ataatcattt tatcatgaaa 780
ataagcagct caagagaatt tttgaagaat cttttatttc ttttaacattt aaccacatga 840
atTTTTaatt tTTTTTgca atacatttaa accgaaatgg tcaaacgac aaccaactga 900
tctttattct aataaacttc tagtttacat ttgcatgtga gtgcatcatc attatcatat 960
ttgtacacaa caaacaagaa aaaaatataa acaatatttt atttaaatat ttatattcca 1020
ctttgactgt agatattaaa tcttgtcatc atttatagtc tcaatattat aattTTTTtta 1080
TTTTTTcaaa attcaaaagt ttacaattat tTTTTgaac tataatatta tccaagatga 1140
acatctcaag aagaaaatta ttaatattgt tatggTTaaa atTTTtacata caatacttgt 1200
TTTTTgcttt actTTTatct taccgtagat acacaatcga cgataactta gtgatcacac 1260
aataataatt atTTTgttca tgacacaata tttataagaa atacttattt ctttctTTta 1320
tccttcagta gttcataata aaaacatacc ataatatttg tgatgcattc atagtacgta 1380

30

atgaaatgac aatttatgtc aaattatttt cttttatact ctcaaacctc ccgtaaaggt 1440
gagatgagtc atttatccaa ttatacataa atatgtcttt attcatgctc tttatcacat 1500
tctgacacat tcacttaatt tcaagagtaa gcaagcatga taactgaaac tatttatgcg 1560
tatcttacct tgatatttga cacattacat gacacacctc aacatcactt tcaaagatta 1620
agcgcaccac catattatct ttcttttttt ttttatgaag gttttataaa attattaaat 1680
taggtccaaa aaattgtttg tcaaataacc ttttatacta gattgatgac aaaaattacc 1740
tttacgtttt gaaagaccat tttaagacct aatctatcag tgactcctta aagttggcac 1800
aatatttcac ttagacaccc taattgaatg atgttcattt taaacaccca atgtagggtt 1860
ccgctatata attttgacac atttcttaac atcaacaaaa atatataatg agtatgtgat 1920
atactcgga atgacgtgaa aaatgaagac atttggttatt tgtatcaaag tagttactaa 1980
ataattaatt ttgaataaaa ataaaagctg accagtaaat caataacaca taatattttc 2040
cacctaataa ttaaaatata aaataaaaaa gagccatctc agggtcattc gccaccatt 2100
gctatttcaa agaaatttgt acgttagttt atagaaattg atgttaaaat tctttcaaga 2160
aaaatttatg aatgaattta ttctctaatt taaaaatatt ttctgttatt tttgttgaaa 2220
gaaatttaac ttggataaaa tgggtggttaa aactggaaag aagaaaagag aaaaaataat 2280
taaaaatcat ttcacgctct aatcaatgag cgtatcacat tcattatggt atataagcaa 2340
aagtgacaaa acgaaaataa tatattacat gaaatgtcta aaataaatat cgtctaatta 2400
aaatatctaa gtaacatatt gtgcctaact ttagagggat catcaataag ttaaacccca 2460
ttttaataac tcataattgt cctttttatt taatattgtc acaaatcaca atgataatta 2520
acattaattt gtcctttgtg acgtccatat tcatgcattt aaccaatcat cttcatttgg 2580
acttattatc acaattatcc cactttcctc acaaaatgga gcattcaagt ggaatagact 2640
acacgatttt taatttcac aaaaacatct ttttgcttta ttcattatta tattgtcgct 2700
attgttgaat tttatttgcc cttaaatttct taccataaat agatttttct tttagaaaaa 2760
ggagattgac taattctttt cttgtaggaa aaggtttagg actctataaa tagagacata 2820
ttccttctaa cttaatcaac atttacaatg tagtcttaaa gactttgaaa gtttttggtt 2880
agggggagaa attgtgggtc acaagcttga tacgttatca attgtgtaaa cctcccatgt 2940
attctgagtg aatttggttg aggttgtttc cctctgtatt ttgtactctc atatttatag 3000
tggattgttc atctctttcg tggacgtagg tcgattgacc gtcgattgac cgaaccacgt 3060
taaactcttg tattttttga tatatttctc attatcttct tactcgtgat ctttcaaggt 3120

ttgcattgct atcttccgcg ttacaccaac ttatttacga tcctaacagc tatgggtgtgg 3180
aaacataaat caaacatttt actgatataa acacatcttt gattataaca tgatagaaat 3240
ttgagcccaa ctttttatca tcattatata caaaaagttc taaatttttt ttttgatgta 3300
gtaaaactta aatccatagt cttgccctta aaccaatgac ataatatata acccaaaata 3360
tactagtttt cgccctcgag ccctttaaaa agtatagtca atatttacgg tgaccgtgaa 3420
tttcttaatt atgatataa atttaaaaga aatcatgac acattctact gatgagaaca 3480
tgtgctaate aagggaaaac atggatgtga aaaatacttt ttgttaaaag taaaaaaaaa 3540
tgtgaaattt tgtagttat ttactaccta tacattattt gagcatgtgc aaactttaca 3600
aatacctaate agaagatttt cacctgcctg tatatatgta aattaattat aatgaacact 3660
ctcacataaa ataattatca gtatatacat taatacttgc cctccacaat gaattaaata 3720
aatgtagaa catgatctac acttcaataa aactaagacc ataaagaata atttcaaaat 3780
atacacatgt caacaataaa ttatttgcac attatattaa cttactaaac aatctttact 3840
tttgaaatat aaaaataate aagttataag tctgctcaaa gtaaagcact tgtagactc, 3900
atctgatttt gagaaggtaa gcaaatgat ggtgcataat agtcacaagt aaaatataaa 3960
atagatttca ttagtaaaat tgttttttac tttctttata tataattatc aatatccttc 4020
aatggtaggt taattatatt gttaacttct tgtagaatta aagcaataag acaagaatat 4080
taaagataaa agaacaataa aaatagaaag actaagagat aagagttttc ttattcttct 4140
ttcaataagt atcatcaagt gtatacaata taaatttttg tatttttgat ctatctattt 4200
ataatgttat atataagcat acaaaagatc agtcataaat atgactttaa tcatgaaaat 4260
aatgaaagag attatgaagg cgtaagggtta ctagaataat agtcattaaa aaaaggggtt 4320
atctttataa ttgaataatt gatgaagtaa tggagataat tagtgagcat aaattttttt 4380
aaaaaatgg acattttacac tataatattt tataacactt tcccttaaac atctaggtat 4440
aaataatgag tcttgtcaaa atcttagtag gaaaaattct gtgaaatttt ttagtgaaa 4500
acaaatgata taaatatctt gaatactcat tatttggtgt ctcattaaaa atcttatctg 4560
acctataaaa taaattattt gctcaactca aaatagtttt tcattctaaa attagtataa 4620
ttattagtga atatttaatt aacataattg tatactaagg ggcctataaa ttggattctt 4680
ctcaaagaaa aataaaatca ccacacaact ttcttcttct gtcacatcaat tagcaattaa 4740
tccaaaacca ttatggctgc 4760

<210> 19
<211> 1229
<212> DNA
<213> Tomato

<220>
<221> promoter
<222> (1)..(1229)
<223>

<400> 19
gatcttactt taccataatg gtgaaaagga tagagaccca catgggtttt acttcgttat 60
agagacaaga tgaaaacaaa tctaaaattt aatattatag atggatagat gatggacaac 120
aaaaagagaa aagaagatac tggtcattgg tccaaaacag ccacccgaat caatatatga 180
ccgaaaaaca aaagctacag aatcatatct gtgcaacggg gccacagtgc tataggatag 240
cacaaccaca ctgtcacata aaaaagagga ttttgcactc gttttagatg gagtttcgta 300
attttcgggt ctttcaagct taaatatata cttcattaaa gcttcgaatt ttgtaatggt 360
caattctacc tctttgatgt tcgataccta taaaataatt aaataaacgt atagacgtag 420
gaacaattaa gcggagttag atagtgcatt tatgattcta cctgtgagtg caatggtaaa 480
atggacatta taaaagagta ggggcaaaga gggaagtga aaattctccc cacttagcca 540
tgtttaatat agtagggata ggaatatgta ataagtagtg ttttttctat ttaattttct 600
gtatacttct tccatctcct ttaattatta aaaggttttc ctctctttac tctttctctc 660
taaattacta ttctgaagta tattttcttt tataaaaaga gtaataaact ttatttccat 720
taaaagaaca aacaacaaga aatgataatc aaatacacat tcatattttt aaaaaaaag 780
ttaaacaaga tatagaaata gttatcaaat atatttatgt tgtcattcct tgtatacaat 840
ggcattcctt tagctttggt tatgtatttc ctgagcttct cttagtgtac tatatccttt 900
aatattaatg catctttcga tcttgctaag atatgataaa aatagacgac acgtgtcaca 960
acctaattga gatatttcga tgtactttct atccgtctta gcttgtaatt aattattggt 1020
aaaaagaat actcaattaa ctagaaacaa gaaataagaa acgaaaacat taaaaacgg 1080
agttgaagcg tgcaaatttg tggaaatgat tgttatcatg aaccagaaaa cattaaataa 1140
ctcttcctat aaaaggcctt tattcttcac tttctcaa atcacgtcctaa agatatcaaa 1200
gatttcaact gatagcaaaa agcactact 1229

<210> 20

33

<211> 845
<212> DNA
<213> Tomate

<220>
<221> promoter
<222> (1)..(845)
<223>

<400> 20
ctgttattga atttctataa aatgttataa tattgatttc ttaatgatca gttaactacg 60
tgattatttg atatgttttt aatctaaaat gtgatatgta aaatatagaa gaaaaaaaaat 120
taaaaagaac ttttaagaaaa aaatttcaac ccacccaac ctaaaatcct aggtccgcca 180
tggttaattat agatatatga tgatgaaggg caaatattgg tctatgagaa tttcgggtgat 240
actaccgctt gaagagcaat aatgggttttg ggactccgat gagggaaaca ttcaaatatg 300
atggattttg gtgatactat gtttaccoga gctagctatc acagaataat ctacatccca 360
caaatgaaat atgttatagg ctaccaatta ggaagtagtg gaattatgaa gaagtaggga 420
tgtgcaaata taagagaaaa tttgaaaatt atgattgaaa caagttatgt ttttttaact 480
agatgaatta aatgggtttta agatttgtag atttataatc aaacaattac cgctactcta 540
tcggtgacta ccaattccat cattgtaaat aacaaataac agattcgttg ctggatgtct 600
tagtgccgtg aagcctacaa atcacactat aaactgctta gctctcgagc gttactaatt 660
tggtgattac caattccaac attgcgactt cttctactag tagtactaaa atagcaagta 720
atatgcattt gtggtaagat gtttggtggt aacctttcct aaccagacta taaatgacct 780
caacactata gtggagtttc atcgatcatc attctaaacg aaaaacttga agtgaaagca 840
tcaag 845

<210> 21
<211> 3417
<212> DNA
<213> Tomate

<220>
<221> promoter
<222> (1)..(3417)
<223>

<400> 21
aagcttggct gcaggtcgac ctgcagggtca acggatcaat gccttggtta taatatgaaa 60
ataagacgta aaagaagtct tgcatatgca ccataatatt agacttatgg acaaaagtaa 120

gttggttcaa attacgcttt tatttatcca catagcaaga aaataatact caaaatccaa 180
cggatcgggt tattttatat tttactctac atgtatatat gtagtataat ggacataaat 240
tctgtcgtaa ttatacatat attaataatg aggattgtaa aataatatgc aaaaacgtcg 300
tatttgacat actaatagct aaaatactac ctactatcat atataattag ttaactatgt 360
gccttttaag aaaaattacg tgaaataaca aatattttaga gcatattatg taatatagct 420
gtagttttat tattttttgt taatggctac aatttcgcaa aattttccta ttttgtttct 480
taatcgata aatccaaatt ttgtataatt atgacctaa ttgtttaatt cagatttcgt 540
ataaaattcg atttttgatt ttataaatta aaatttatac ttacttttagc tacttgttta 600
tgatttatca aaaaattcat attaattctat ttgtatatgg acaagcaaaa tatacaaattg 660
gagttctgaa aatttctaaa tgcataact taatatcttt gatggtcact caactatcaa 720
ctttttccat aaaaagtcac ttaacattga ttttcaactc gaaaatcact caactatgaa 780
atctttgtat agaaagtcac tcaacctatt taattatfff tttccattat atctgttgtc 840
acgaaatatt atttctaact aatattctaa gaataaacat acatccattt aaatcattta 900
ataaaccgc ccacttgacc taaccacat aatattaaca cttttgtttt actttttatc 960
tccaaaatta ttttcttgggt ttccattct tttctctttg cttttttttt cttcttctca 1020
atttcagcct ttttcttct ttttttagta aacctcagtc aaataggaat tagattgtga 1080
ttaaataatt attagaagga tgcagggttg tacaaagaga gtttattaag agataatcta 1140
taaaaaaaaa aaagtcagat atgcatatt cagattcaga gatcattaaa tgatgacttt 1200
tttcgtaata ggttttcttt aaatcctttc gccttcatac gacgactctc gataataaca 1260
tcgttttaag ctaataatgc taatgaacaa taatcaaat aaaaaagaat tcggatacaa 1320
gagaaaatga tttagtga gaaaaaattg agatattcct tattcctaac taaacgaagg 1380
aagaagaggc taaaattgag attcagttaa aaaaaaaaaa caaagaaaaa cgcaatggag 1440
atgagagaaa gtaattttga aaaataaaaa taaattaaga gggtaaataat tttattttta 1500
gcgagttggg ttaagtgggt cgggtcatta aatggatata tgttttatttc ttaaaatttt 1560
agttagaaat acaaatttca aatcaacaaa ttttaatgaa aaaataatta aatagggtga 1620
gtggctttct atgcaaagat ctcatagttg agtgattttt gagtagaaaa tcatagttaa 1680
gtgagtttct gtgaaaaaaaa attgatagtt gagtgactat caaagatatt aactctagac 1740
ttgtcatatt cgtatactta catacgaaat atacaaacct ctgcctccat gacaagcaaa 1800

35

aaactataac tatgaaacaa tatttttcgaa atcatagcta taaagtctta ttatatctaa 1860
tatctttact atttttaaaa atttcacata attttaatac ataaataatt tactttttaac 1920
taacgaaaaa ggacattttt atgtcacctg agagcccatc ggtagattca tcacattttt 1980
tsgtttcttg taataaactg tacacatata aggagaaatt aaattagaga ttatttttcc 2040
attttgagga gattaataaa tttaaaatgt aacttaacat gtaaactgct ataaaggtaa 2100
caaacacgt aaactgctat aaaggtaatt ctattttaaaa gataaataaa tgcttaaaaag 2160
aagtgccaaa aaaacacaaa caaacaatg aaactaaacc tacttcaagg gaagttcttg 2220
tagtataaaa ataaataaag tcaacttatt cacgacattt ctttttggtt ttcttttggc 2280
tacgtattca tatttaagtc tgactaattt agattctcgc tatatataaa agattcaggg 2340
gtggctcaac gcaattggag gcctagagca aaatttcaat tcgcggccta atatattata 2400
tactttatat acctatttat tcaaaattta ttttttttac actatttaga tggaaattat 2460
tagtacttaa tattgttttt tcagttatta gtttaggta aaattttatt aatacaacat 2520
tgaaaaacat cttttaagtg agacaattat tatatgtatt gttaacatag tgctataagt 2580
aataagtaaa taaatattaa ataaaaataa gagtaagaac catagaattt gacacaagaa 2640
gttgatgact tggatatact cattttaaca tgcttgact ttagtaatgc ttgaatctaa 2700
aattttaaaaa gaaataaaaa agaatttgta atccactttt tccaacactt ttactgtta 2760
attcttattt ttaacatagt acaaaaaata ttaaatgga taaaataatt tattttataa 2820
aagattatat atatattttt ttatcatata taactaattt ttctataaaa atttaaacac 2880
ataatttaat tttaaaaaaa atttggggct ttggggccta agacaaaggc cttaaaggac 2940
aaaacataga gccgccctg aaaagatctc attcgaaaga aaatatgcat taccaatgat 3000
ttttcgtacc cagagctcaa aatcaaaatt gtactgttat ttttttaaaa aatttcatct 3060
cagactaaat ggaatttttt tctttgggta acctgtttga tcaatctttt ggaatcagtt 3120
aattttgaaa aataaattaa tgagaaataa tttgtatttg tccagcttat ttaagaatta 3180
tttttgagca acaatttata ttagtcacg cttttaagtg tattttttta aataaaatta 3240
aggtattatt tgaaaaaatt acttttaaaa aaattgaatt aaattctgtt actcttatta 3300
tatactccta tataatttga ttgccaaaaa tatcaaactg ttaatatattg aagttgatgt 3360
gagggattac ttcttgatta aattgtacta caatgtaata ttatcaaatt aaagctt 3417

<210> 22

<211> 1155

36

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (6)..(995)

<223>

<400> 22

gaagc atg cag cta gca gcg aca gta atg ttg gag cag ctt acc gga agc 50
 Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser

1

5

10

15

gct gag gca ctc aag gag aag gag aag gag gtt gca ggc agc tct gac 98
 Ala Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp
 20 25 30

gtg ttg cgt aca tgg gcg acc cag tac tcg ctt ccg tca gag gag tca 146
 Val Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser
 35 40 45

gac gcg gcc cgc ccg gga ctg aag aat gcc tac aag cca cca cct tcc 194
 Asp Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser
 50 55 60

gac aca aag ggc atc aca atg gcg cta gct gtc atc ggc tcc tgg gcc 242
 Asp Thr Lys Gly Ile Thr Met Ala Leu Ala Val Ile Gly Ser Trp Ala
 65 70 75

gca gtg ttc ctc cac gcc att ttt caa atc aag ctt ccg acc tcc ttg 290
 Ala Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu
 80 85 90 95

gac cag ctg cac tgg ctg ccc gtg tca gat gcc aca gct cag ctg gtt 338
 Asp Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val
 100 105 110

agc ggc agc agc agc ctg ctg cac atc gtc gta gta ttc ttt gtc ctg 386
 Ser Gly Ser Ser Ser Leu Leu His Ile Val Val Val Phe Phe Val Leu
 115 120 125

gag ttc ctg tac aca ggc ctt ttt atc acc acg cat gat gct atg cat 434
 Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His
 130 135 140

ggc acc atc gcc atg aga aac agg cag ctt aat gac ttc ttg ggc aga 482
 Gly Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg
 145 150 155

gta tgc atc tcc ttg tac gcc tgg ttt gat tac aac atg ctg cac cgc 530
 Val Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg
 160 165 170 175

aag cat tgg gag cac cac aac cac act ggc gag gtg ggc aag gac cct 578
 Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro
 180 185 190

37

gac ttc cac agg gga aac cct ggc att gtg ccc tgg ttt gcc agc ttc 626
 Asp Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe
 195 200 205

atg tcc agc tac atg tcg atg tgg cag ttt gcg cgc ctc gca tgg tgg 674
 Met Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp
 210 215 220

acg gtg gtc atg cag ctg ctg ggt gcg cca atg gcg aac ctg ctg gtg 722
 Thr Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val
 225 230 235

ttc atg gcg gcc gcg ccc atc ctg tcc gcc ttc cgc ttg ttc tac ttt 770
 Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe
 240 245 250 255

ggc acg tac atg ccc cac aag cct gag cct ggc gcc gcg tca ggc tct 818
 Gly Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser
 260 265 270

tca cca gcc gtc atg aac tgg tgg aag tcg cgc act agc cag gcg tcc 866
 Ser Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser
 275 280 285

gac ctg gtc agc ttt ctg acc tgc tac cac ttc gac ctg cac tgg gag 914
 Asp Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu
 290 295 300

cac cac cgc tgg ccc ttt gcc ccc tgg tgg gag ctg ccc aac tgc cgc 962
 His His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg
 305 310 315

cgc ctg tct ggc cga ggt ctg gtt cct gcc tag ctggacacac tgcagtgggc 1015
 Arg Leu Ser Gly Arg Gly Leu Val Pro Ala
 320 325

cctgctgcca gctgggcatg caggttgtgg caggactggg tgaggtgaaa agctgcaggc 1075
 gctgctgccg gacacgctgc atgggctacc ctgtgtagct gccgccacta ggggaggggg 1135
 tttgtagctg tcgagcttgc 1155

<210> 23
 <211> 329
 <212> PRT
 <213> Haematococcus pluvialis

<400> 23
 Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala
 1 5 10 15
 Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val
 20 25 30

38

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp
35 40 45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp
50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Ala Val Ile Gly Ser Trp Ala Ala
65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp
85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser
100 105 110

Gly Ser Ser Ser Leu Leu His Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr
210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
245 250 255

39

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser
 260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp
 275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His
 290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg
 305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala
 325

<210> 24
 <211> 1111
 <212> DNA
 <213> Haematococcus pluvialis

<220>
 <221> CDS
 <222> (4)..(951)
 <223>

<400> 24
 tgc atg cta gag gca ctc aag gag aag gag aag gag gtt gca ggc agc 48
 Met Leu Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser
 1 5 10 15

tct gac gtg ttg cgt aca tgg gcg acc cag tac tcg ctt ccg tca gaa 96
 Ser Asp Val Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu
 20 25 30

gag tca gac gcg gcc cgc ccg gga ctg aag aat gcc tac aag cca cca 144
 Glu Ser Asp Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro
 35 40 45

cct tcc gac aca aag ggc atc aca atg gcg cta gct gtc atc ggc tcc 192
 Pro Ser Asp Thr Lys Gly Ile Thr Met Ala Leu Ala Val Ile Gly Ser
 50 55 60

tgg gcc gca gtg ttc ctc cac gcc att ttt caa atc aag ctt ccg acc 240
 Trp Ala Ala Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr
 65 70 75

tcc ttg gac cag ctg cac tgg ctg ccc gtg tca gat gcc aca gct cag 288
 Ser Leu Asp Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln
 80 85 90 95

40

ctg gtt agc ggc agc agc agc ctg ctg cac atc gtc gta gta ttc ttt Leu Val Ser Gly Ser Ser Ser Leu Leu His Ile Val Val Val Phe Phe 100 105 110	336
gtc ctg gag ttc ctg tac aca ggc ctt ttt atc acc acg cat gat gct Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala 115 120 125	384
atg cat ggc acc atc gcc atg aga aac agg cag ctt aat gac ttc ttg Met His Gly Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu 130 135 140	432
ggc aga gta tgc atc tcc ttg tac gcc tgg ttt gat tac aac atg ctg Gly Arg Val Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu 145 150 155	480
cac cgc aag cat tgg gag cac cac aac cac act ggc gag gtg ggc aag His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys 160 165 170 175	528
gac cct gac ttc cac agg gga aac cct ggc att gtg ccc tgg ttt gcc Asp Pro Asp Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala 180 185 190	576
agc ttc atg tcc agc tac atg tcg atg tgg cag ttt gcg cgc ctc gca Ser Phe Met Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala 195 200 205	624
tgg tgg acg gtg gtc atg cag ctg ctg ggt gcg cca atg gcg aac ctg Trp Trp Thr Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu 210 215 220	672
ctg gtg ttc atg gcg gcc gcg ccc atc ctg tcc gcc ttc cgc ttg ttc Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe 225 230 235	720
tac ttt ggc acg tac atg ccc cac aag cct gag cct ggc gcc gcg tca Tyr Phe Gly Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser 240 245 250 255	768
ggc tct tca cca gcc gtc atg aac tgg tgg aag tcg cgc act agc cag Gly Ser Ser Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln 260 265 270	816
gcg tcc gac ctg gtc agc ttt ctg acc tgc tac cac ttc gac ctg cac Ala Ser Asp Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His 275 280 285	864
tgg gag cac cac cgc tgg ccc ttc gcc ccc tgg tgg gag ctg ccc aac Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn 290 295 300	912
tgc cgc cgc ctg tct ggc cga ggt ctg gtt cct gcc tag ctggacacac Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala 305 310 315	961
tgcaagtgggc cctgctgccca gctggggcatg caggttggtgg caggactggg tgaggtgaaa	1021

41

agctgcaggc gctgctgccg gacacgttgc atgggctacc ctgtgtagct gccgccacta 1081
ggggagggggg tttgtagctg tcgagcttgc 1111

<210> 25
<211> 315
<212> PRT
<213> Haematococcus pluvialis

<400> 25

Met Leu Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser
1 5 10 15

Asp Val Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu
20 25 30

Ser Asp Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro
35 40 45

Ser Asp Thr Lys Gly Ile Thr Met Ala Leu Ala Val Ile Gly Ser Trp
50 55 60

Ala Ala Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser
65 70 75 80

Leu Asp Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu
85 90 95

Val Ser Gly Ser Ser Ser Leu Leu His Ile Val Val Val Phe Phe Val
100 105 110

Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met
115 120 125

His Gly Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly
130 135 140

Arg Val Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His
145 150 155 160

Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp
165 170 175

42

Pro Asp Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser
 180 185 190

Phe Met Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp
 195 200 205

Trp Thr Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu
 210 215 220

Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr
 225 230 235 240

Phe Gly Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly
 245 250 255

Ser Ser Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala
 260 265 270

Ser Asp Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp
 275 280 285

Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys
 290 295 300

Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala
 305 310 315

<210> 26
 <211> 1031
 <212> DNA
 <213> Haematococcus pluvialis

<220>
 <221> CDS
 <222> (6)..(1031)
 <223>

<400> 26
 gaagc atg cag cta gca gcg aca gta atg ttg gag cag ctt acc gga agc 50
 Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser
 1 5 10 15

gct gag gca ctc aag gag aag gag aag gag gtt gca ggc agc tct gac 98
 Ala Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp
 20 25 30

gtg ttg cgt aca tgg gcg acc cag tac tcg ctt ccg tca gag gag tca 146

43

Val	Leu	Arg	Thr	Trp	Ala	Thr	Gln	Tyr	Ser	Leu	Pro	Ser	Glu	Glu	Ser		
			35					40					45				
gac	gcg	gcc	cg	ccg	gga	ctg	aag	aat	gcc	tac	aag	cca	cca	cct	tcc		194
Asp	Ala	Ala	Arg	Pro	Gly	Leu	Lys	Asn	Ala	Tyr	Lys	Pro	Pro	Pro	Ser		
		50					55					60					
gac	aca	aag	ggc	atc	aca	atg	gcg	cta	gct	gtc	atc	ggc	tcc	tgg	gct		242
Asp	Thr	Lys	Gly	Ile	Thr	Met	Ala	Leu	Ala	Val	Ile	Gly	Ser	Trp	Ala		
	65					70				75							
gca	gtg	ttc	ctc	cac	gcc	att	ttt	caa	atc	aag	ctt	ccg	acc	tcc	ttg		290
Ala	Val	Phe	Leu	His	Ala	Ile	Phe	Gln	Ile	Lys	Leu	Pro	Thr	Ser	Leu		
80					85					90					95		
gac	cag	ctg	cac	tgg	ctg	ccc	gtg	tca	gat	gcc	aca	gct	cag	ctg	gtt		338
Asp	Gln	Leu	His	Trp	Leu	Pro	Val	Ser	Asp	Ala	Thr	Ala	Gln	Leu	Val		
				100					105					110			
agc	ggc	agc	agc	agc	ctg	ctg	cac	atc	gtc	gta	gta	ttc	ttt	gtc	ctg		386
Ser	Gly	Ser	Ser	Ser	Leu	Leu	His	Ile	Val	Val	Val	Phe	Phe	Val	Leu		
				115				120					125				
gag	ttc	ctg	tac	aca	ggc	ctt	ttt	atc	acc	acg	cat	gat	gct	atg	cat		434
Glu	Phe	Leu	Tyr	Thr	Gly	Leu	Phe	Ile	Thr	Thr	His	Asp	Ala	Met	His		
		130					135					140					
ggc	acc	atc	gcc	atg	aga	aac	agg	cag	ctt	aat	gac	ttc	ttg	ggc	aga		482
Gly	Thr	Ile	Ala	Met	Arg	Asn	Arg	Gln	Leu	Asn	Asp	Phe	Leu	Gly	Arg		
	145					150				155							
gta	tgc	atc	tcc	ttg	tac	gcc	tgg	ttt	gat	tac	aac	atg	ctg	cac	cg		530
Val	Cys	Ile	Ser	Leu	Tyr	Ala	Trp	Phe	Asp	Tyr	Asn	Met	Leu	His	Arg		
160					165				170					175			
aag	cat	tgg	gag	cac	cac	aac	cac	act	ggc	gag	gtg	ggc	aag	gac	cct		578
Lys	His	Trp	Glu	His	His	Asn	His	Thr	Gly	Glu	Val	Gly	Lys	Asp	Pro		
				180					185					190			
gac	ttc	cac	agg	gga	aac	cct	ggc	att	gtg	ccc	tgg	ttt	gcc	agc	ttc		626
Asp	Phe	His	Arg	Gly	Asn	Pro	Gly	Ile	Val	Pro	Trp	Phe	Ala	Ser	Phe		
			195				200						205				
atg	tcc	agc	tac	atg	tcg	atg	tgg	cag	ttt	gcg	cg	ctc	gca	tgg	tgg		674
Met	Ser	Ser	Tyr	Met	Ser	Met	Trp	Gln	Phe	Ala	Arg	Leu	Ala	Trp	Trp		
		210					215					220					
acg	gtg	gtc	atg	cag	ctg	ctg	ggc	gcg	cca	atg	gcg	aac	ctg	ctg	gtg		722
Thr	Val	Val	Met	Gln	Leu	Leu	Gly	Ala	Pro	Met	Ala	Asn	Leu	Leu	Val		
	225					230					235						
ttc	atg	gcg	gcc	gcg	ccc	atc	ctg	tcc	gcc	ttc	cg	ttg	ttc	tac	ttt		770
Phe	Met	Ala	Ala	Ala	Pro	Ile	Leu	Ser	Ala	Phe	Arg	Leu	Phe	Tyr	Phe		
240					245					250					255		
ggc	acg	tac	atg	ccc	cac	aag	cct	gag	cct	ggc	gcc	gcg	tca	ggc	tct		818
Gly	Thr	Tyr	Met	Pro	His	Lys	Pro	Glu	Pro	Gly	Ala	Ala	Ser	Gly	Ser		

44

260

265

270

tca cca gcc gtc atg aac tgg tgg aag tcg cgc act agc cag gcg tcc 866
 Ser Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser
 275 280 285

gac ctg gtc agc ttt ctg acc tgc tac cac ttc gac ctg cac tgg gag 914
 Asp Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu
 290 295 300

cac cac cgc tgg ccc ttt gcc ccc tgg tgg gag ctg ccc aac tgc cgc 962
 His His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg
 305 310 315

cgc ctg tct ggc cga ggt ctg gtt cct gcc gag caa aaa ctc atc tca 1010
 Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Glu Gln Lys Leu Ile Ser
 320 325 330 335

gaa gag gat ctg aat agc tag 1031
 Glu Glu Asp Leu Asn Ser
 340

<210> 27
 <211> 341
 <212> PRT
 <213> Haematococcus pluvialis

<400> 27

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala
 1 5 10 15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val
 20 25 30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp
 35 40 45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp
 50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Ala Val Ile Gly Ser Trp Ala Ala
 65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp
 85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser
 100 105 110

45

Gly Ser Ser Ser Leu Leu His Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr
210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser
260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp
275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His
290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg
305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala Glu Gln Lys Leu Ile Ser Glu
325 330 335

46

Glu Asp Leu Asn Ser
340

<210> 28
<211> 777
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> promoter
<222> (1)..(777)
<223>

<400> 28
gagctcactc actgatttcc attgcttgaa aattgatgat gaactaagat caatccatgt 60
tagtttcaaa acaacagtaa ctgtggccaa cttagttttg aaacaacact aactgggtcga 120
agcaaaaaga aaaaagagtt tcatcatata tctgatttga tggactgttt ggagtttagga 180
ccaaacatta tctacaaaca aagacttttc tctaacttg tgattccttc ttaaacccta 240
ggggtaatat tctattttcc aaggatcttt agttaaggc aaatccggga aattattgta 300
atcatttggg gaaacatata aaagatttga gttagatgga agtgacgatt aatccaaaca 360
tatatatctc tttcttctta tttcccaaata taacagacaa aagtagaata ttggctttta 420
acaccaatat aaaaacttgc ttcacaccta aacacttttg tttacttttag ggtaagtgca 480
aaaagccaac caaatccacc tgcactgatt tgacgtttac aaacgccgtt aagtcgatgt 540
ccgttgattt aaacagtgtc ttgtaattaa aaaaatcagt ttacataaat ggaaaattta 600
tcacttagtt ttcatcaact tctgaactta cttttcatgg attaggcaat actttccatt 660
tttagtaact caagtggacc ctttacttct tcaactccat ctctctctt ctatttcact 720
tctttcttct cattatatct cttgtcctct ccaccaaata tcttcaacaa aaagctt 777

<210> 29
<211> 22
<212> DNA
<213> kuenstlich

<220>
<221> primer_bind
<222> (1)..(22)
<223>

<400> 29
gcaagctcga cagctacaaa cc

47

<210> 30
<211> 24
<212> DNA
<213> kuenstlich

<220>
<221> primer_bind
<222> (1)..(24)
<223>

<400> 30
gaagcatgca gctagcagcg acag

24

<210> 31
<211> 30
<212> DNA
<213> kuenstlich

<220>
<221> primer_bind
<222> (1)..(30)
<223>

<400> 31
tgcattgctag aggcactcaa ggagaaggag

30

<210> 32
<211> 59
<212> DNA
<213> kuenstlich

<220>
<221> primer_bind
<222> (1)..(59)
<223>

<400> 32
ctagctattc agatcctctt ctgagatgag tttttgctcg gcaggaacca gacctcggc

59

<210> 33
<211> 28
<212> DNA
<213> kuenstlich

<220>
<221> primer_bind
<222> (1)..(28)
<223>

48

<400> 33
gagctcactc actgatttcc attgcttg

28

<210> 34
<211> 37
<212> DNA
<213> kuenstlich

<220>
<221> primer_bind
<222> (1)..(37)
<223>

<400> 34
cgccgttaag tcgatgtccg ttgatttaaa cagtgtc

37

<210> 35
<211> 34
<212> DNA
<213> kuenstlich

<220>
<221> primer_bind
<222> (1)..(34)
<223>

<400> 35
atcaacggac atcgacttaa cggcgtttgt aaac

34

<210> 36
<211> 25
<212> DNA
<213> kuenstlich

<220>
<221> primer_bind
<222> (1)..(25)
<223>

<400> 36
taagcttttt gttgaagaga tttgg

25